



UNIVERSITAT<sub>DE</sub>  
BARCELONA

## ***Insulators: estructura y funciones***

Ujué Fresán Salvo



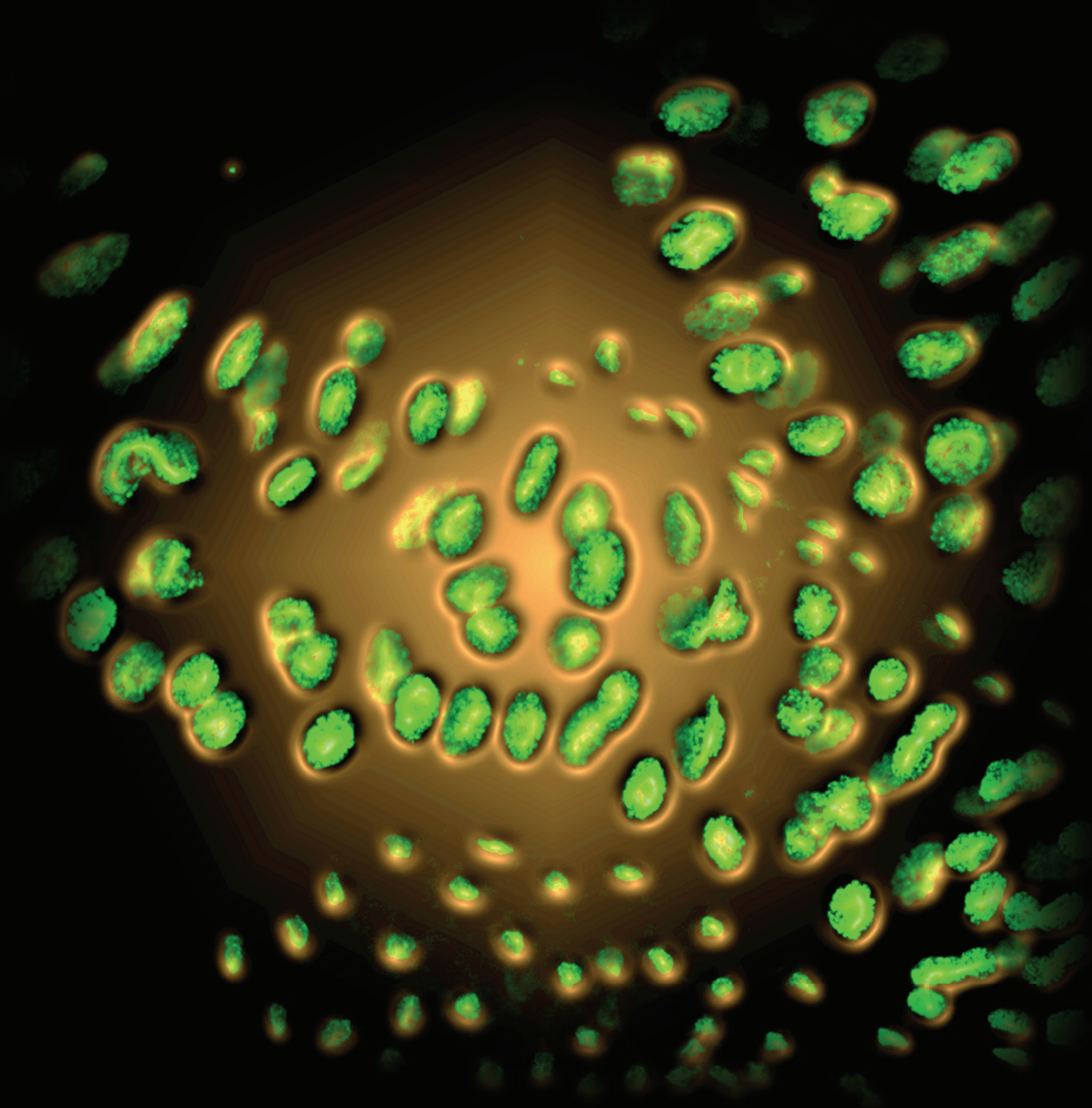
Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 3.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 3.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 3.0. Spain License.**

TESIS DOCTORAL

*Insulators:*  
estructura y funciones



Ujué Fresán Salvo









UNIVERSITAT DE BARCELONA  
FACULTAD DE FARMACIA  
DOCTORADO EN BIOMEDICINA

## ***INSULATORS: estructura y funciones***

Memoria presentada por Ujué Fresán Salvo para optar al título de doctor por la Universitat de Barcelona. Este trabajo ha sido realizado bajo la dirección de la Dra. Maria Lluïsa Espinàs Janer en el Instituto de Biología Molecular de Barcelona (IBMB-CSIC) y el *Institut de Recerca Biomèdica* (IRB Barcelona).

Vistobueno de la directora

Vistobueno de la tutora

Dra. Maria Lluïsa Espinàs Janer

Dra. Maribel Geli Fernández-Peñaflor

Autor

Ujué Fresán Salvo

Barcelona, Marzo 2016



A mis padres y a Carlos,  
por su apoyo incondicional



# AGRADECIMIENTOS

Parece que nunca iba a llegar este día, pero todo tiene un principio y un final. Durante estos años he crecido no solo profesionalmente sino también como persona, con gente tan maravillosa que es imposible describirla. Os voy a echar mucho de menos.

En primer lugar me gustaría dar las gracias a Lluïsa y a todo su grupo. Tras unas prácticas de verano confiaste en mí para comenzar con el proyecto de Haspin, y aquí está el resultado. Por supuesto, Sergi, sin ti no hubiera sido posible ni tan siquiera comenzar. Ahora que yo también sé qué es el estar acabando una tesis te agradezco profundamente todo el tiempo que me has dedicado, y encima siempre con una sonrisa. María, ahora te quedas tú al mando del carro (no el de Manolo Escobar que te estoy viendo...). Sigue con esa alegría que te caracteriza, que los clonajes y las inmunos, si es con buen humor, se llevan mejor.

Ali, qué te voy a decir a ti. Te lo he recalcado una y mil veces. Has sido un apoyo fundamental en esta etapa. Te deseo todo lo mejor en tu nueva vida, y espero que me permitas compartirla contigo. Por supuesto Olga, mi gurú. Lo sabes todo (excepto hacer politénicos y ChIPs...). Gracias por cada uno de tus valiosos consejos, no sé qué hubiera sido de mi tesis sin ti. Sonieta, tu sonrisa y alegría hacía que mereciera la pena venir cada día a trabajar, a pesar de que Haspin y BAF hayan sido un poco rebeldes. Laura, nuestra historia ha sido breve pero intensa. He descubierto en ti una magnífica persona y una gran amiga, y espero que esto siga así para siempre. Evidentemente que no me olvido de Salva, con tus míticas frases y locuras que tanto me han alegrado el día a día, además de todos los conciertos y teatros a los que hemos ido juntos. Albert, gracias por tu buen humor a pesar de que tires los Yatekomos por la cafetería, te toquen los kiwis verdes o te vayas corriendo con media casa y una guitarra a cuestas porque pierdes el autobús. Sabes que soy la fan número uno de Nuar. Eva, my Guardian Angel, one of the craziest girls that I have ever met! Don't ever change! Y en general a todas y cada una de las personas con las que he compartido mi vida en el PBA3: Paula, Anne, Estefanía, Roman, Anna, Carles, Mònica, Tomás, Andrea, Aina, Eva Mejía, Carolina... no continuo la lista porque es enorme, pero GRACIAS a todos y cada uno de vosotros.

Pero no solo a la gente del día a día. Papá, Mamá: a pesar de la distancia os he sentido junto a mí siempre, y vuestro apoyo y el orgullo que sentís por mí me ha dado fuerzas para continuar día a día con la tesis (evidentemente, los tupperes también han ayudado...).

Y finalmente como no podía ser de otra manera, Carlos. Por levantarme cuando me caigo, por entenderme cuando ni yo misma sé qué quiero, por confiar en mí cuando ni yo me creo... Eres el mejor compañero de viaje, no solo en el de vacaciones, sino en el de la vida.

**¡¡GRACIAS!!**



# Índice





## ÍNDICE DE FIGURAS

## ÍNDICE DE TABLAS

## ABREVIATURAS

## INTRODUCCIÓN

1.	ORGANIZACIÓN GENÓMICA.....	1
1.1.	Estructura de la cromatina.....	1
1.1.1.	La cromatina está organizada en dominios.....	2
1.1.1.1.	<i>Chromosome Conformation Capture</i> (3C).....	2
1.1.1.2.	<i>Topologically Associating Domains</i> (TADs) .....	4
1.1.1.3.	Los TADs están separados por fronteras.....	5
1.1.1.4.	Los TADS son unidades funcionales .....	8
1.1.1.5.	Relación entre TADs y rasgos epigenómicos.....	9
	<i>Modificación posttranscripcional de histonas</i> .....	11
	<i>Interacción de la cromatina con la lamina</i> .....	13
1.2.	Regulación de la expresión génica .....	15
1.2.1.	Elementos básicos de la transcripción .....	15
1.2.2.	Mecanismos implicados en la regulación de la transcripción génica .....	16
1.2.2.1.	<i>Enhancers</i> .....	16
1.2.2.2.	Elementos de respuesta a Polycomb .....	19
1.2.2.3.	<i>Insulators</i> .....	22
	CTCF.....	24
	CP190.....	25
	Cohesinas .....	27
	Condensinas .....	29
2.	DESARROLLO DE <i>DROSOPHILA MELANOGASTER</i> .....	30
2.1.	Regulación de genes homeóticos.....	31
2.1.1.	Genes homeóticos.....	31
2.1.2.	El complejo Bithorax .....	31
2.1.2.1.	Organización y funcionamiento de la región reguladora de <i>Abdominal-B</i> .....	32
2.2.	Regulación temporal de la esteroidogénesis .....	34

## **Índice**

2.2.1.	Transiciones entre estadíos larvarios y metamorfosis .....	34
2.2.2.	Genes <i>Halloween</i> .....	35
3.	HASPIN.....	36
3.1.	Haspin y la segregación del material genético .....	38
3.1.1.	Mitosis .....	38
3.1.1.1.	Dinámica de unión de cohesinas y condensinas al DNA .....	40
3.1.1.2.	Segregación cromosómica en células pluripotenciales .....	42
3.1.2.	Meiosis.....	43
3.1.2.1.	Dinámica de unión de cohesinas y condensinas al DNA .....	43
<b>OBJETIVOS</b> .....		49

## **MATERIALES Y MÉTODOS**

1.	MATERIALES .....	53
1.1.	Oligonucleótidos.....	53
1.2.	Plásmidos.....	55
1.3.	Anticuerpos .....	56
1.4.	Líneas de <i>Drosophila</i> .....	57
1.5.	Cepas bacterianas.....	59
1.6.	Líneas celulares.....	59
2.	MÉTODOS .....	60
2.1.	Mantenimiento de cultivos celulares .....	60
2.2.	Técnicas de manipulación de <i>D. melanogaster</i> .....	60
2.2.1.	Ensayo <i>enhancer-blocking</i> .....	60
2.2.2.	Cuantificación de la coloración del pigmento de ojos.....	61
2.2.3.	Sistemas de control de la expresión de genes en <i>Drosophila</i> .....	62
2.2.3.1.	<i>Heat-shock</i> .....	62
2.2.3.2.	Sistema <i>UAS-Gal4</i> .....	62
2.2.4.	Estudio de viabilidad.....	63
2.2.5.	Duración de la vida de moscas adultas .....	63
2.2.6.	Estudio del tiempo de desarrollo .....	63
2.3.	Técnicas de trabajo con DNA.....	63

2.3.1.	Reacción en cadena de la polimerasa (PCR) .....	63
2.3.2.	Transformación de células competentes .....	64
2.3.3.	Minipreparación de DNA plasmídico .....	64
2.3.4.	Maxipreparación de DNA plasmídico .....	65
2.3.5.	Extracción de DNA genómico de mosca .....	65
2.3.6.	Extracción de DNA genómico de glándulas salivares .....	65
2.3.7.	Cuantificación del tamaño de los núcleos de glándulas salivares .....	66
2.4.	Técnicas de trabajo con RNA .....	66
2.4.1.	Extracción de RNA .....	66
2.4.2.	Retrotranscripción y PCR cuantitativa .....	67
2.4.	Técnicas de trabajo con proteínas .....	68
2.4.1.	Extracción proteica .....	68
2.4.2.	Electroforesis en gel de acrilamida .....	68
2.4.3.	<i>Western Blot</i> .....	69
2.4.4.	Obtención y purificación de anticuerpos .....	69
2.4.4.1.	Expresión y purificación de proteínas fusionadas a una cola de histidinas .....	69
	<i>Pruebas de expresión</i> .....	70
	<i>Expresión a gran escala</i> .....	70
	<i>Purificación a partir de la fracción insoluble</i> .....	71
2.4.4.2.	Obtención de anticuerpos policlonales .....	73
	<i>Purificación de anticuerpos por afinidad</i> .....	74
2.4.5.	Immunotinciones .....	74
2.4.5.1.	Cromosomas politénicos .....	74
2.4.5.2.	<i>Squash</i> de cerebros .....	75
2.4.5.3.	Discos imaginales de ala y glándulas salivares .....	76
2.4.6.	Co-Immunoprecipitación .....	76
2.4.7.	Immunoprecipitación de cromatina (ChIP) .....	77
2.4.7.1.	Secuenciación y análisis .....	78
2.5.	Técnicas de trabajo con lípidos .....	80
2.5.1.	Tinción con <i>Oil Red O</i> .....	80

### RESULTADOS

1. NUEVAS FUNCIONES DE PROTEINAS <i>INSULATORS</i> .....	83
1.1. CTCF regula el desarrollo en <i>D. melanogaster</i> .....	83
1.2. CTCF actúa en la vía de síntesis de ecdisona regulando la expresión de los genes <i>Halloween</i> .....	84
1.3. Esta función de CTCF es independiente de otras proteínas <i>insulators</i> .....	88
1.4. CTCF es necesaria para la homeostasis del colesterol y síntesis de ecdisona.....	89
2. CARACTERIZACIÓN DE PROTEINAS QUE INTERERACTÚAN CON LA PROTEÍNA <i>INSULATOR</i>	
CP190.....	91
2.1. Análisis funcional de Ibf1 e Ibf2.....	91
2.1.1. Caracterización de Ibf1 e Ibf2.....	91
2.1.2. Ibf1 e Ibf2 son proteínas con actividad <i>enhancer-blocking</i> .....	91
2.1.3. Ibf1 e Ibf2 forman <i>insulator bodies</i> , y se necesitan mutuamente.....	93
2.2. Caracterización y análisis funcional de pita.....	95
2.2.1. Características de pita .....	95
2.2.2. Pita forma parte del complejo CP190-Ibf1-Ibf2.....	95
2.2.3. Pita tiene actividad <i>enhancer-blocking</i> .....	96
3. NUEVAS PROTEINAS CON ACTIVIDAD <i>INSULATOR</i> : Haspin.....	97
3.1. Identificación de Haspin como proteína <i>insulator</i> .....	97
3.2. Haspin .....	97
3.3. Fenotipos .....	99
3.4. Haspin es una proteína con actividad <i>enhancer-blocking</i> .....	100
3.5. Actividad kinasa .....	102
3.5.1. Mitosis .....	102
3.5.2. Interfase.....	103
3.6. Organización nuclear .....	104
3.6.1. Localización de Haspin en interfase .....	104
3.6.2. La localización de CP190 y la estructura nuclear están alteradas en ausencia de Haspin.....	105
3.6.3. La ausencia de Haspin afecta a la condensación de la cromatina.....	107
3.6.3.1. Haspin provoca cambios en el tamaño de los núcleos en interfase .....	107
3.6.3.2. Haspin no afecta a la replicación.....	109

3.6.3.3.	El fenotipo es producido gracias al dominio kinasa .....	109
3.7.	Análisis de la localización de H3T3ph en el genoma .....	111
3.8.	Relación de Haspin con otras proteínas involucradas en la organización de la cromatina .....	115
3.8.1.	Cohesinas .....	115
3.8.1.1.	La unión de pds5 al DNA depende de Haspin .....	115
3.8.1.2.	Pds5 regula la condensación de la cromatina .....	118
3.8.1.3.	Pds5 y cohesinas tienen actividad <i>enhancer-blocking</i> .....	119
3.8.2.	Condensinas .....	120
3.8.2.1.	La presencia nuclear de CapH2 no depende de Haspin .....	120
3.8.2.2.	Los efectos de Haspin sobre la condensación de la cromatina dependen de la presencia de CapH2.....	121
3.8.2.3.	Los sitios enriquecidos en H3T3ph no tienen CapH2, y viceversa .....	122

## **DISCUSIÓN**

1.	NUEVAS FUNCIONES DE PROTEÍNAS <i>INSULATORS</i> .....	125
1.1.	CTCF regula el tiempo de desarrollo en <i>Drosophila</i> .....	125
1.2.	CTCF es necesaria para la homeostasis del colesterol .....	128
2.	PROTEÍNAS QUE INTERACTÚAN CON CP190 .....	129
2.1.	Ibf1 e Ibf2 son proteínas <i>insulators</i> .....	130
2.1.1.	Ibf1 e Ibf2 tienen actividad <i>enhancer-blocking</i> .....	130
2.1.2.	Ibf1 e Ibf2 se necesitan mutuamente para localizar en los <i>insulator bodies</i> .....	132
2.2.	Pita como <i>insulator</i> .....	135
3.	NUEVAS PROTEÍNAS CON ACTIVIDAD <i>INSULATOR</i> : Haspin .....	137
3.1.	Haspin tiene actividad <i>enhancer-blocking</i> .....	138
3.2.	Haspin es necesaria para la organización de la cromatina en el núcleo en interfase.....	138
3.3.	Haspin ejerce su acción a través de su dominio kinasa .....	140
3.4.	Mecanismo de acción de Haspin.....	143

<b>CONCLUSIONES</b> .....	153
---------------------------	-----

<b>BIBLIOGRAFÍA</b> .....	157
---------------------------	-----

## **Índice**

---

<b>ANEXO I</b> .....	181
----------------------	-----

<b>ANEXO II</b> .....	189
-----------------------	-----

## **INTRODUCCIÓN**

Figura 1. Los niveles de empaquetamiento de la cromatina.....	1
Figura 2. <i>Chromatin conformation capture</i> da información sobre la organización 3D del genoma.....	3
Figura 3. Los TADs están separados por fronteras.....	5
Figura 4. Las proteínas <i>insulators</i> determinan la fuerza de la frontera entre TADs .....	7
Figura 5. Características de los 5 tipos de cromatina.....	10
Figura 6. La cromatina se divide en función de 9 tipos de estados. ....	11
Figura 7. Modificaciones covalentes de las histonas. ....	13
Figura 8. Modelo de organización de la cromatina en la periferia nuclear. ....	14
Figura 9. Activación de la transcripción génica .....	18
Figura 10. <i>Polycomb response element</i> .....	21
Figura 11. Funciones clásicas de <i>insulators</i> .....	22
Figura 12. Complejo de las cohesinas .....	27
Figura 13. Complejo de las condensinas I y II.....	29
Figura 14. El complejo Bithorax y los segmentos que controla. ....	32
Figura 15. Modelo de los Dominios.....	33
Figura 16. Niveles de ecdisona durante el desarrollo. ....	34
Figura 17. Localización de Aurora B en <i>inner</i> centrómeros depende de Haspin y Bub1. ....	39
Figura 18. Regulación de la segregación mitótica.....	41
Figura 19. H3T3ph es necesaria para la segregación asimétrica del material genético.....	42
Figura 20. Dinámica de unión de cohesinas y condensinas. ....	45



## Índice de figuras

### MATERIALES Y MÉTODOS

Figura 21. Ensayo <i>enhancer-blocking</i> . .....	61
Figura 22. Sistema <i>UAS-Gal4</i> para expresar o inhibir un gen.....	62
Figura 23. El extremo N-terminal de Haspin se expresa de forma insoluble. ....	71
Figura 24. Purificación de la proteína de fusión Haspin-His.....	72

### RESULTADOS

Figura 25. La ausencia de CTCF produce retraso en el desarrollo.....	83
Figura 26. CTCF regula los niveles de expresión de algunos, pero no todos, de los genes que intervienen en la vía de señalización de la ecdisona.....	85
Figura 27. CTCF regula la expresión de los genes <i>Halloween</i> en la glándula protorácica.....	86
Figura 28. CTCF no afecta a la expresión de los genes <i>PTTH</i> ni <i>Torso</i> . ....	87
Figura 29. Retraso en el desarrollo de larvas que no expresan <i>CTCF</i> exclusivamente en glándula protorácica. ....	87
Figura 30. Niveles transcripcionales de los genes <i>Halloween</i> . ....	88
Figura 31. Tiempo de desarrollo de larvas que no expresan <i>CP190</i> o <i>Ibf1</i> e <i>Ibf2</i> en glándula protorácica.. ....	89
Figura 32. Ambos, 20-hidroxiecdisona (20E) y colesterol (C) son necesarios para recuperar el fenotipo que produce la ausencia de CTCF en la glándula protorácica.....	89
Figura 33. Acumulación de lípidos en la glándula protorácica, en larvas control ( <i>phm&gt;w<sup>-</sup></i> ) y <i>phm&gt;CTCF-RNAi</i> . ....	90
Figura 34. <i>Ibf1</i> e <i>Ibf2</i> tienen actividad <i>enhancer-blocking</i> en <i>Fab8</i> . ....	92
Figura 35. <i>Ibf1</i> e <i>Ibf2</i> forman <i>insulator bodies</i> y colocalizan con <i>CP190</i> . ....	93
Figura 36. <i>Ibf1</i> e <i>Ibf2</i> se necesitan para localizar en <i>insulator bodies</i> , pero <i>CP190</i> no depende de ellas. ....	94
Figura 37. Pita coimmunoprecipita con <i>CP190</i> , <i>Ibf1</i> e <i>Ibf2</i> . ....	96

Figura 38. Pita es necesaria para la actividad <i>enhancer-blocking</i> en Fab6 .....	96
Figura 39. Disminución del nivel de expresión del gen <i>Haspin</i> en la línea <i>Haspin</i> <sup>86</sup> .....	98
Figura 40. Generación de un mutante nulo para <i>Haspin</i> . .....	99
Figura 41. Disminución del nivel de expresión del gen <i>Haspin</i> en la línea <i>Haspin</i> <sup>128</sup> .....	99
Figura 42. La ausencia de Haspin no causa letalidad, pero sí afecta a la supervivencia. ....	100
Figura 43. Haspin es necesaria para la actividad <i>enhancer-blocking</i> en diferentes Fabs del complejo Bithorax. ....	101
Figura 44. Haspin tiene actividad <i>enhancer-blocking</i> . ....	102
Figura 45. Haspin fosforila H3T3 en la región pericentromérica. ....	103
Figura 46. H3T3ph está presente en interfase. ....	104
Figura 47. Haspin se localiza en la periferia nuclear. ....	105
Figura 48. Haspin afecta a la localización nuclear de CP190 sin interferir con su unión al DNA. ....	106
Figura 49. Haspin no afecta a la localización nuclear de otras proteínas <i>insulators</i> .....	106
Figura 50. Haspin afecta a la estructura nuclear. ....	107
Figura 51. Haspin afecta al tamaño nuclear. ....	108
Figura 52. La mutación de <i>Haspin</i> no provoca problemas de replicación del DNA en glándulas salivares. ....	109
Figura 53. Haspin afecta a la condensación de la cromatina a través de su dominio kinasa. ....	110
Figura 54. Haspin regula la condensación nuclear a través de su dominio kinasa. ....	111
Figura 55. H3T3ph se localiza preferente, pero no exclusivamente, en regiones pericentroméricas. ....	112
Figura 56. Comparación de la unión de H3T3ph con otras proteínas. ....	112
Figura 57. H3T3ph está ausente en zonas activas .....	113
Figura 58. H3T3ph se une a la heterocromatina pericentromérica y en las regiones silenciadas. ....	113
Figura 59. H3T3ph se une a genes homeóticos. ....	114

## **Índice de figuras**

Figura 60. Polycomb disminuye su unión a los Fabs del complejo Bithorax en ausencia de Haspin..	114
Figura 61. Pds5 es una proteína nuclear que se une al DNA.....	115
Figura 62. Haspin afecta a la unión de pds5 al DNA.....	116
Figura 63. La cantidad de pds5 es menor en ausencia de Haspin. ....	116
Figura 64. La cantidad de pds5 y su unión al DNA dependen de Haspin .....	117
Figura 65. El nivel de expresión de <i>pds5</i> no depende de Haspin .....	117
Figura 66. Pds5 regula la condensación de la cromatina .....	118
Figura 67. Haspin y pds5 regulan la condensación nuclear.....	119
Figura 68. Pds5 tiene actividad <i>enhancer-blocking</i> . ....	119
Figura 69. SMC1 tiene actividad <i>enhancer-blocking</i> .. ....	120
Figura 70. La presencia nuclear de CapH2 no depende de Haspin.. ....	120
Figura 71. CapH2 regula la condensación de la cromatina. ....	121
Figura 72. El efecto en la condensación que produce la falta de Haspin requiere de la presencia de CapH2 .....	122
Figura 73. Perfil de unión de H3T3ph y CapH2 a la cromatina.....	122

**MATERIALES Y MÉTODOS**

Tabla 1. Lista de oligonucleótidos utilizados.....	54
Tabla 2. Lista de plásmidos utilizados .....	55
Tabla 3. Lista de anticuerpos primarios utilizados .....	56
Tabla 4. Lista de anticuerpos secundarios utilizados. ....	56
Tabla 5. Líneas de <i>Drosophila melanogaster</i> utilizadas .....	59



## ABREVIATURAS

Abreviatura	Significado
<i>abd-A</i>	Gen <i>abdominal-A</i>
<i>Abd-B</i>	Gen <i>Abdominal-B</i>
Ant-C	Complejo Antennapedia
AEL	Después de la puesta de los huevos
BEAF	<i>Boundary element-associated factor of 32kDa</i>
cDNA	DNA monocatenario
ChIP	Immunoprecipitación de la cromatina
ChIP-seq	Secuenciación de la inmunoprecipitación de la cromatina
Co-IP	Coimmunoprecipitación
CPC	<i>Cromosomal Passenger Complex</i>
CP190	<i>Centrosomal Protein 190</i>
CTCF	<i>CCCTC-binding factor</i>
CTD	Dominio C-terminal de la enzima RNA polimerasa II
DAPI	4'-6-diamidino-2-fenilindole
DCC	<i>Dosage compensation complex</i>
<i>dib</i>	Gen <i>disembodied</i>
DNA	Ácido desoxirribonucleico
<i>EcR</i>	Receptor de ecdisona
Fab	Dominio Front-abdominal
GAF	Factor GAGA
GFP	<i>Green Fluorescence Protein</i>
GSC	Células pluripotenciales de línea germinal
GTF	Factores generales de transcripción
Haspin	<i>Haploid germ cell-specific nuclear protein kinase</i>
Hi-C	<i>Genome-wide Chromosome Conformation Capture</i>
Hox	Gen homeótico
iab	Dominio infra-abdominal
lbf	<i>Insulator binding factor</i>
ICA	eje entre cromátidas
IP	Immunoprecipitación
Kb	Kilobase
LCR	<i>Locus Control Region</i>
Mod(mdg4)	<i>Modifier of mdg4</i>
modENCODE	<i>MODEL organisms ENCYclopedia Of DNA Elements</i>
mRNA	RNA mensajero
<i>nobo</i>	Gen <i>noppera-bo</i>
NPC	Complejo de poros nucleares
<i>nvd</i>	Gen <i>neverland</i>
pb	Pares de bases

## **Abreviaturas**

<i>Pc</i>	Gen <i>Polycomb</i>
PcG	Grupo Polycomb
PCR	Reacción en cadena de la polimerasa
pds5	<i>Precocious dissociation of sisters protein 5</i>
<i>phm</i>	Gen <i>phamton</i>
PHO-RC	<i>Pho-repressive complex</i>
PIC	Complejo de preiniciación
PG	Glándula protorácica
PRC	<i>Polycomb Repressive Complex</i>
PRE	<i>Polycomb Response Elements</i>
PTM	Modificación posttranscripcional
PTS	<i>Promoter targeting sequence</i>
PTTH	<i>Protoracicotropic hormone</i>
RNA	Ácido ribonucleico
RNAi	RNA de interferencia
RNApol	RNA polimerasa
rpm	Revoluciones por minuto
SAC	<i>Spindle assembly checkpoint</i>
<i>Sad</i>	Gen <i>shadow</i>
SEM	<i>Error estándar de la media</i>
SMC	<i>Structural maintenance complex</i>
<i>spok</i>	Gen <i>spookier</i>
<i>sro</i>	Gen <i>shroud</i>
Su(Hw)	<i>Supressor of Hairy wing</i>
TAD	<i>Topologically Association Domain</i>
TF	Factor de transcripción
trxG	Grupo trithorax
TSS	Sitio de inicio de la transcripción
UAS	<i>Upstream Activating Sequence</i>
<i>Ubx</i>	Gen <i>Ultrabithorax</i>
UTR	Región no traducida
WT	<i>Wild-type</i>
20E	20-hidroxiecdisona
3C	<i>Chromatin conformation capture</i>

# Introducción





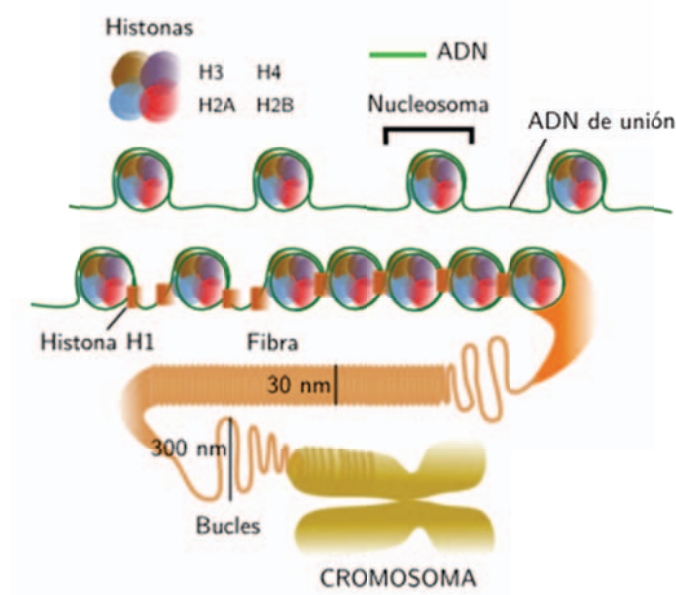
## 1. ORGANIZACIÓN GENÓMICA

### 1.1. Estructura de la cromatina

Los organismos eucariotas tienen el DNA (molécula que contiene la información genética necesaria para el desarrollo y mantenimiento de los seres vivos) organizado dentro del núcleo en una estructura denominada cromatina (Figura 1).

La unidad básica de la cromatina es el nucleosoma, que está compuesto por un octámero que incluye dos copias de las histonas H3, H4, H2A y H2B, proteínas que están altamente conservadas. Alrededor de este octámero se asocian unos 147pb de DNA formando el *core* del nucleosoma (Luger et al., 1997). Al DNA que queda entre nucleosomas adyacentes se le denomina DNA de unión.

El conjunto de nucleosomas conectados, con apariencia de collar de perlas, forma una fibra de 11nm de diámetro. Una quinta histona, histona H1, interacciona con el DNA de unión y se une al DNA que entra y sale del nucleosoma, promoviendo y estabilizando la compactación de la fibra nucleosomal en una estructura repetitiva de 30nm de diámetro (Robinson and Rhodes, 2006).



**Figura 1. Los niveles de empaquetamiento de la cromatina.** El DNA se compacta en diferentes niveles. En un primer nivel, alrededor de los octámeros de histonas, formando los nucleosomas. Estos se disponen formando la fibra de 30nm, que se superenrolla alcanzando la estructura terciaria de 300nm. A través de loops y otras superestructuras alcanza una compactación máxima, como se observa en los cromosomas metafásicos.

Más allá de esta estructura secundaria, la cromatina se puede organizar también de manera terciaria mediante interacciones a larga distancia. Estas interacciones, que pueden tomar forma de

## **Introducción**

*loops* (bucles), como los que se producen por las interacciones entre elementos reguladores de cromosomas interfásicos tal que las realizadas por *enhancers* e *insulators*, u otras superestructuras, como las que se encuentran en cromosomas metafásicos (Woodcock and Ghosh, 2010).

Durante muchos años se ha considerado la estructura de la cromatina únicamente como una mera forma de empaquetamiento del material genético, diferenciándose por completo de los mecanismos que controlan la transcripción. Con el paso del tiempo la imagen de la cromatina ha ido cambiando, pasando a ser considerada una estructura dinámica con funciones importantes en procesos fundamentales. De hecho, la distribución espacial y la compartimentalización del genoma tiene gran importancia en procesos como la replicación, reparación y recombinación, pero sobre todo han sido estudiadas en la regulación de la expresión de genes (Sexton et al., 2007).

### **1.1.1. La cromatina está organizada en dominios**

Como hemos dicho, la organización de la cromatina dentro del núcleo está relacionada con la regulación de la expresión de genes. A grandes rasgos, a nivel nuclear se pueden observar diferencias entre la cromatina menos condensada (abierta o activa), y la más condensada (compacta o silenciada). Así, en los cromosomas interfásicos se puede distinguir la cromatina abierta, denominada eucromatina, donde se encuentran la mayor parte de los genes transcripcionalmente activos, y la condensada, la heterocromatina, en forma de bloques de cromatina compacta e inaccesible. De manera que se establecen dos grandes regiones en el genoma, las transcripcionalmente activas, que en general se localizan internamente en el núcleo, y las inactivas, que tienden más hacia la periferia nuclear así como alrededor del nucleolo. Además, cabe señalar que hay regiones eucromáticas que se silencian que se denominan heterocromatina facultativa, en contraposición a la constitutiva, que es la que está siempre silenciada. Además, la disposición de los cromosomas dentro del núcleo no es al azar, y se distribuyen en lo que se llaman territorios cromosómicos (Cremer and Cremer, 2010). Inicialmente se habló de que los más pequeños se situaban de manera más central, mientras que los más grandes tenderían hacia la periferia. Años más tarde se observó que los que son ricos en genes se encuentran en una posición central, mientras que los que no tienen gran cantidad tienden hacia la periferia, independientemente de su tamaño (Boyle et al., 2001).

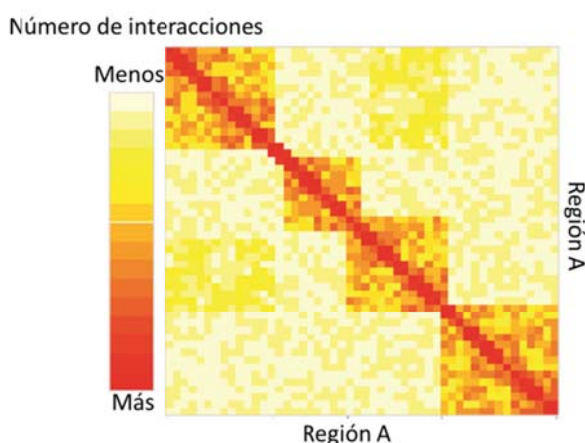
#### **1.1.1.1. Chromosome Conformation Capture (3C)**

Dekker (Dekker et al., 2002) desarrolló una estrategia bioquímica para analizar la frecuencia de contactos entre diferentes sitios del genoma: *Chromosome conformation capture* (3C). Este método, y los que posteriormente se basaron en él, permiten realizar un estudio topológico del DNA más sistemático a alta resolución que la microscopía. Se ha conseguido estudiar la organización de los cromosomas más detalladamente, así como las interacciones a larga distancia que se creían que

ocurrían. Para conseguir su propósito, el de conocer la frecuencia de contactos, se fija la cromatina de manera que las zonas que interactúan se mantengan juntas, y tras ser digerida con enzimas de restricción, los extremos de los fragmentos se ligan intramolecularmente. De esta manera, fragmentos que están muy alejados en la secuencia lineal, pero que colocalizan en el espacio, pueden ser detectados juntos. Se determina la conformación 3D de un gen o de un cromosoma midiendo el número de ligaciones, y por tanto interacciones, entre diferentes zonas del genoma. Posteriormente, a partir de esta técnica, se han ido desarrollando otras, 4C, 5C, ChIA-PET y Hi-C. Esta última permite conocer las interacciones que ocurren entre todas las zonas del genoma.

Aunque inicialmente se aplicó en levadura (Dekker et al., 2002), pronto se adaptó a mamíferos, demostrando que existen *loops* de cromatina *in vivo* entre elementos reguladores del DNA y sus genes diana, tal como el gen de  $\beta$ -globina. Se observó que el *Locus Control Region* (LCR), a gran distancia del gen que regula, interactuaba físicamente con los genes de globina activos, formando de esta manera un *loop* que dejaba fuera la cromatina que quedaba entre ambos (Tolhuis et al., 2002).

Los resultados de las interacciones de diferentes zonas de la cromatina obtenidos con estas técnicas se muestran en un gráfico de dos dimensiones, donde en cada uno de los ejes se representa la región analizada (Figura 2). Normalmente poseen una diagonal prominente que refleja los contactos frecuentes entre sitios localizados juntos en la secuencia lineal. El análisis de estos mapas llevó a la primera observación de que existen interacciones entre secuencias dentro de un dominio, que se representan como unos cuadros de relativa alta frecuencia de interacción alrededor de la diagonal. Estos dominios que interaccionan consigo mismos están separados por fronteras muy marcadas que parece que estructuralmente les aíslan de los dominios adyacentes, ya que con el contiguo presenta mucha menor interacción, a pesar de la proximidad física.



**Figura 2.** *Chromatin conformation capture* da información sobre la organización 3D del genoma. Representación de las interacciones a lo largo de una hipotética región genómica. La graduación de coloración indica la frecuencia de interacción entre diferentes zonas. Se observa que la región está dividida en dominios con muchas interacciones dentro del mismo, aunque también se muestran algunas interacciones entre secuencias de diferentes dominios (Modificado de Graaf and van Steensel, 2013).

## **Introducción**

Su interpretación lleva a concluir que la cromatina está dividida en dominios cuyas secuencias interaccionan con ellas mismas. Dentro de estos dominios hay una coasociación entre genes separados en la cromatina. Se ha visto también que existen interacciones entre secuencias de distintos cromosomas, en que los contactos ocurren a baja frecuencia, pero mucho mayor que la esperada por azar. Se dan generalmente entre aquellos que comparten regulación por algún factor común, como los que están reprimidos por proteínas del grupo de Polycomb (PcG), o su activación en un tejido específico. Algunos grupos incluso hablan de redes espaciales de funcionamiento de genes, donde el reclutamiento de genes a un lugar del núcleo enriquecido en sus factores reguladores facilitaría la coordinación de su expresión (Sexton and Cavalli, 2015).

### **1.1.1.2. Topologically Associating Domains (TADs)**

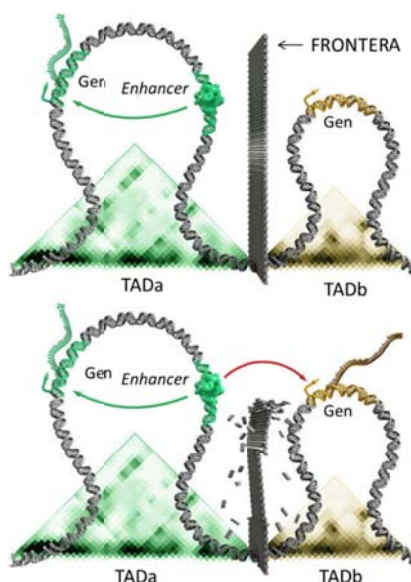
Estos dominios que acabamos de describir, caracterizados por la alta probabilidad de contacto entre las secuencias dentro de él mismo según se deriva de los resultados de las técnicas 3C, están presentes en una gran cantidad de especies, como *Drosophila* y mamíferos, y se denominan *Topologically Associating Domains* (Nora et al., 2012). Mientras que la organización de la cromatina en TADs se ve en diferentes organismos, su mecanismo de formación y los complejos de proteínas involucrados, su organización interna y sus funciones en procesos de regulación genómica como en transcripción y replicación varían mucho.

En el caso concreto de *D. melanogaster* su genoma está dividido en 1000 dominios de entre 10 y 100 Kb cada uno, equivalente a un TAD. Fue inicialmente estudiado por el grupo de Cavalli (Sexton et al., 2012) y por el de Corcés (Hou et al., 2012). Estos dominios se correlacionan con rasgos epigenómicos tal como modificaciones de histonas, densidad de genes activos, interacción con estructuras nucleares, tiempo de replicación o la composición en nucleótidos u elementos repetitivos.

En mamíferos se vio que la disrupción de grandes dominios silenciados en el *X chromosome inactivation centre* (*Xic*) no afectaba a la segmentación de los TADs, aunque sí a la compactación local (Nora et al., 2012). Además, comparando diferentes líneas celulares y tejidos humanos y de ratón se vio que muchos dominios permanecían invariables (Dixon et al., 2012). Esto sugería que los TADs eran una configuración espacial donde los rasgos regulatorios se establecían *a posteriori*. Sin embargo, estudiando genes homeóticos de ratón se vio que a medida que se iban silenciando unos genes y activando otros, la configuración en TADs cambiaba (Noordermeer et al., 2011). La investigación de la diferenciación celular de células ES llevó a concluir que la realidad es una mezcla entre ambas posturas, ya que los TADs son predominantemente estables pero pueden desarrollar pequeños “sub-TADs” transitorios (Phillips-Cremins et al., 2013).

### 1.1.1.3. Los TADs están separados por fronteras

La siguiente pregunta sería cómo las interacciones entre secuencias ocurren dentro del mismo dominio y no con otros contiguos, y cómo se forman estos TADs. Se ha propuesto que deben existir fronteras muy marcadas entre los TADs con propiedades de aislamiento, que facilitan el contacto dentro del mismo, y dificultan la comunicación con otros, a pesar de la proximidad espacial (Sexton et al., 2012) (Figura 3), como se ha derivado de varios estudios. En uno se produjo la eliminación de una frontera entre los genes *Xist-Tsix* en células ESC de ratón, promoviendo interacciones aberrantes entre TADs que previamente estaban separados y, en consecuencia, la desregulación de genes, probablemente por nuevas interacciones *enhancer*-promotor. Esta eliminación originó una fusión parcial entre los TADs previamente separados. Sin embargo, los dominios no se fusionaron completamente, lo que indicaba que otros factores además de los de la frontera, de dentro de los propios TADs, también están involucrados en mantener esta separación cuando no existe tal frontera (Nora et al., 2012).



**Figura 3. TADs están separados por fronteras.** Representación de la activación de un gen dentro de un TAD por su *enhancer*, que no puede actuar sobre otro gen situado en un TAD diferente separado por una frontera (panel superior). Cuando se altera la frontera, se pueden establecer conexiones *enhancer*-promotor entre regiones que estaban localizadas en dominios diferentes antes de la ruptura de la misma (panel inferior) (Modificado de Molecular Genetics/ Thomas Splettstoesser, [www.scistyle.com](http://www.scistyle.com)).

Se ha visto que la separación entre dominios se produce por una combinación entre transcripción activa, alta densidad génica y presencia de proteínas *insulators* (Hou et al., 2012). No solo las proteínas clásicamente consideradas como tal, sino también otros factores arquitectónicos se encuentran en estos bordes, como cohesinas (Rad21), subunidades de TFIIC o condensinas (Caph2), y se ha correlacionado la cantidad de proteínas arquitectónicas situadas en la frontera con la fuerza de la misma (Figura 4A) (Van Bortle et al., 2014, Li et al., 2015).

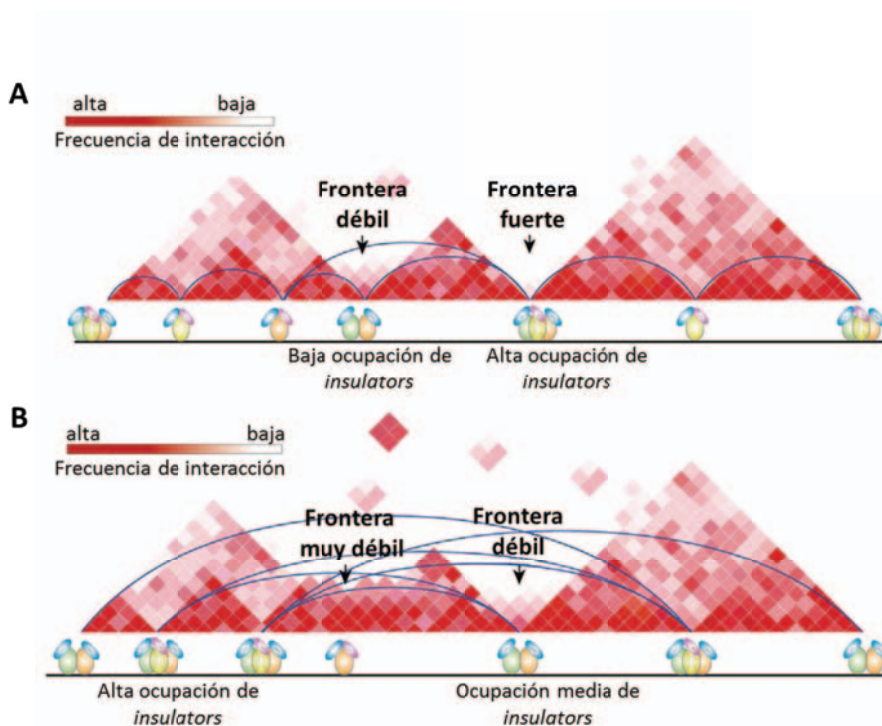
## **Introducción**

Las proteínas *insulators* son fundamentales en las fronteras en diferentes organismos. Un estudio reciente realizado sobre cuatro especies de mamíferos, ha revelado mediante Hi-C que la organización modular de los cromosomas está fuertemente conservada, y que se correlaciona con el mantenimiento de los sitios de unión de CTCF (Vietri Rudan et al., 2015), que son justamente aquellos en los que colocaliza con cohesinas, y están enriquecidos en las fronteras de los TADs. Además, las diferencias vistas entre especies en los sitios de unión de CTCF se correlacionan con diferencias en la estructura interna de los dominios. De hecho, aunque son necesarios mapas de interacción de más organismos para sacar conclusiones, se ha visto que las especies que conservan la organización en TADs son aquellas que también conservan la proteína CTCF (Heger et al., 2012, Sexton and Cavalli, 2015).

No obstante, en mamíferos solo el 15% de las secuencias de CTCF y de cohesinas se encuentran en fronteras, mientras que el 85% restante se localiza dentro de los TADs. Cómo desempeñan su función de borde entre dominios varía entre diferentes experimentos. Mientras que la depleción de cohesinas en células HEK293 produce una pérdida general de interacciones intracromosomales sin afectar a la organización del TAD, la falta de CTCF causa una disminución similar en la frecuencia de interacciones dentro de un TAD con un aumento de las que se producen entre TADs adyacentes (Zuin et al., 2014). En cambio, la carencia de cohesinas en astrocitos postmitóticos de ratón afecta a las interacciones intraTADs mediadas por ellas y CTCF, produciendo también un debilitamiento de la frontera entre TADs (Sofueva et al., 2013).

En *Drosophila* las fronteras entre TADs, como en mamíferos, dependen, al menos en parte, de proteínas *insulators* (Figura 4A) (Hou et al., 2012). Un estudio ha demostrado que los cambios en la organización espacial de los cromosomas que se producen tras un choque térmico están relacionados con la afectación de estas proteínas (Li et al., 2015). Este estrés supone una reorganización 3D, y las proteínas *insulators* que normalmente se encuentran en las fronteras pasan a localizarse dentro de los TADs (Figura 4B). Este cambio provoca un aumento de interacciones a larga distancia entre TADs, con un incremento de los contactos entre *enhancers* y promotores silenciados por proteínas Polycomb.

Al contrario que en mamíferos, son numerosas las proteínas de este tipo descritas en este animal (CTCF, BEAF-32, Su(Hw)...) (Hou et al., 2012, Sexton et al., 2012) y cada una reconoce un motivo de DNA (Cubenas-Potts and Corces, 2015). Se ha visto que éstas interaccionan con múltiples proteínas accesorias (CP190, Rad21, CapH2...), por lo que las combinaciones entre estas dos clases de proteínas, así como el número y orientación de los sitios de unión producen una gran variedad de posibilidades de organización 3D.



**Figura 4. Las proteínas *insulators* determinan la fuerza de la frontera entre TADs.** **A/** Representación de tres dominios diferentes en los que se ve la posición de las proteínas *insulators* así como la fuerza de la frontera en función de la cantidad de proteínas que tiene. Los semicírculos azules muestran algunos de los contactos establecidos. La graduación de coloración indica la frecuencia de interacción entre diferentes zonas. **B/** Representación de los cambios que sufre la estructura de los TADs y los contactos entre ellos respecto de la figura A tras un choque térmico (Modificado de Li et al., 2015).

Aunque es evidente el papel de las proteínas *insulators* en las fronteras, cómo organizan la cromatina para prevenir las interacciones entre secuencias pertenecientes a diferentes TADs no está del todo claro. El aislamiento puede llevarse a cabo a través de una actividad local del complejo proteico unido a una frontera, actuando como barrera en la comunicación de elementos reguladores presentes en diferentes dominios. Otra posibilidad es que las proteínas arquitectónicas ejerzan su acción a través de interacciones entre fronteras, por ejemplo mediante la unión de CTCF y/o cohesinas entre ellos (Dowen et al., 2014, Rao et al., 2014) se promovería una configuración en *loop*. Esta estructura física aislaría a los genes localizados en ella respecto a los que están fuera de la misma (Doyle et al., 2014), impidiendo que un *enhancer* localizado en un TAD active un promotor que se encuentra en otro diferente. De hecho, se ha visto que las interacciones a larga distancia entre fronteras son incluso mayores que las que se dan entre la secuencia de un mismo dominio (Hou et al., 2012), y que se producen más eficientemente cuando los sitios de unión de CTCF tienen una orientación invertida (Ali et al., 2016).

Sin embargo, aún queda mucho por descubrir sobre la formación de los TADs. La estructura global de los cromosomas está regulada a lo largo del ciclo celular y se ha visto que los TADs, conservados



## **Introducción**

durante la interfase, se pierden en la mitosis, para posteriormente volverse a formar, pero no se sabe cómo este proceso se realiza ni qué factores pueden estar involucrados (Naumova et al., 2013).

### 1.1.1.4. Los TADS son unidades funcionales

Cada vez hay más evidencias de que los TADs no son solo elementos estructurales, sino que también son dominios funcionales. Se ha visto que la organización espacial y la compartimentalización del genoma repercuten en diversas funciones nucleares, como la transcripción y la replicación.

Los primeros datos que relacionan los TADs con la expresión génica vienen del análisis de la expresión de genes dentro del mismo dominio durante la diferenciación celular en células ES (Nora et al., 2012). Se vio que los genes que estaban situados en el mismo TAD tenían dinámica de expresión similar, y que ésta era diferente a la de los que estaban en otros dominios. Otro estudio posterior probó que los TADs son regulados como unidades funcionales a través de la inducción de la expresión génica por la hormona progestina en células T47D (Le Dily et al., 2014), donde hasta el 20% de estos dominios se comportaban como unidades regulatorias, y la mayoría de los genes que contenían estaban o activados o reprimidos.

La evidencia más fuerte se obtuvo insertando en ratones un gen *reporter* en diferentes localizaciones de un cromosoma, de manera que se vio que los que estaban insertados dentro del mismo dominio tenían el mismo patrón de expresión, lo que implicaba que todos los promotores de los genes de ese dominio estaban activados de igual manera (Symmons et al., 2014).

Recientemente otros estudios han mostrado que la reestructuración que afecta a la organización de un TAD altera la expresión génica por el cambio en los patrones de las interacciones entre *enhancer* y promotor que produce. Una inversión en el cromosoma 3 está implicada en la leucemia mieloide aguda. Esto provoca la disrupción de 2 TADs y lleva a la formación de uno nuevo híbrido (Groschel et al., 2014). Como resultado, el *enhancer* que normalmente regula *GATA2* se coloca en una nueva posición en el TAD, que contiene el oncogén *EVI*. El *enhancer* lo activa y contribuye a la formación del tumor. Al mismo tiempo el gen *GATA2* ya no está colocalizando en el mismo dominio que su *enhancer*, produciendo haploinsuficiencia del *GATA2*, implicado en este tipo de leucemia, y mielodisplasia familiar esporádica, así como en los síndromes MonoMac/Emberger. Éstas y otras observaciones indican que los contactos ectópicos que alteran la expresión de genes pueden producirse por alteración de estas fronteras entre dominios, resaltando su importancia en la compartimentalización genómica y la regulación génica (Lupianez et al., 2015).

Además del papel en la regulación génica, también se ha establecido una relación entre los TADs y la replicación del DNA, aunque el mecanismo de cómo lo regula no se conoce con detalle (Pope et al., 2014). Los tiempos de replicación varían a lo largo de los cromosomas en unidades de varios cientos de kilobases. Se ha visto que los bordes de los dominios de replicación coinciden con las fronteras entre los TADs, al menos en algunos tipos celulares. Por otro lado, una serie de TADs que se

encuentran juntos pueden replicarse todos de manera temprana, pero la transición hacia una replicación tardía ocurrirá en la frontera entre TADs, aunque en cuál ocurra dependerá del tipo celular. De acuerdo con esto, los dominios normalmente se replican como una unidad, de manera temprana o tardía, y activan su replicación como unidades durante la diferenciación en función de los cambios con su actividad transcripcional y el estado de la cromatina.

### 1.1.1.5. Relación entre TADs y rasgos epigenómicos

Los principios de la organización de los cromosomas también se han estudiado mediante el mapaje de los patrones de ocupación a lo largo del genoma de modificaciones de histonas y proteínas de unión a la cromatina.

En el año 2010, Filion hizo un análisis integrativo de los lugares de unión a lo largo de todo el genoma de 53 proteínas de unión a la cromatina, que incluía enzimas que modifican histonas, otras que se unen específicamente a estas modificaciones, componentes de la maquinaria de transcripción, remodeladores de nucleosomas, proteínas con actividad *insulator*, heterocromáticas, componentes estructurales de la cromatina y una selección de factores de unión al DNA, así como modificaciones de histonas, utilizando para ello células de *Drosophila*. Este análisis provee una visión muy completa de la composición del genoma (Filion et al., 2010) (Figura 5). La investigación concluye que el genoma está fragmentado principalmente en 5 tipos de cromatina (haciendo una clasificación por colores: azul, verde, negra, roja y amarilla) que quedan definidos por una combinación de proteínas y forman dominios que pueden llegar a ser incluso mayores de 100kb. Estos tipos de cromatina se corresponden bastante bien con los TADs, asemejándose los cambios en los tipos de cromatina con las fronteras entre los TADs.

La cromatina verde está definida principalmente por la unión de HP1 así como por la presencia de H3K9me2. Es abundante en regiones pericéntricas y en el cromosoma 4.

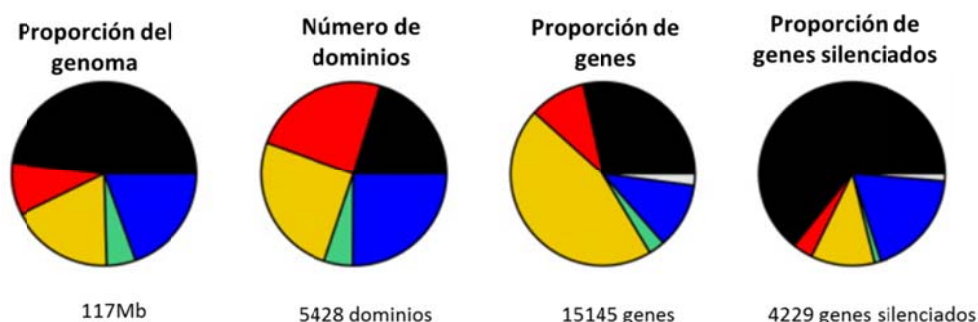
La azul se corresponde con la unida por proteínas del grupo de Polycomb. De hecho, genes que se conocen que son dianas para PcG, como los homeóticos (Hox), se localizan en este tipo. Aunque no se analizó inicialmente, comprueban que la marca producida por este grupo de proteínas, H3K27me3, también está presente en este tipo de cromatina.

La denominada cromatina negra abarca casi la mitad del genoma y, por tanto, es la más abundante. Tiene los dominios más grandes, es pobre en genes, y estos están principalmente silenciados o con poca actividad, como se confirma por la carencia de marcas activadoras como H3K4me2 o H3K79me3, y por su localización preferente en la lamina nuclear.

Las cromatinas roja y amarilla tienen varias características eucromáticas en común, como genes que producen gran cantidad de RNA mensajero (mRNA), altos niveles de RNA polimerasa (RNAPol), H3K4me2 y H3K79me3, y comparten además la carencia de modificaciones como H3K9me2 y

## Introducción

H3K27me3. No obstante, existen también diferencias entre ellas. Los genes de la cromatina roja se replican antes y están implicados en procesos específicos como la unión a receptores, respuesta de defensa, factores de transcripción y transducción de señales. Los de la cromatina amarilla presentan H3K36me3, propia de genes que se expresan constitutivamente, y desempeñan funciones celulares más universales como la reparación del DNA o procesos metabólicos de los ácidos nucleicos.



**Figura 5. Características de los 5 tipos de cromatina.** Relación de cada tipo de cromatina respecto al total en lo que respecta a la proporción que ocupa en el genoma, número de dominios, proporción de genes y proporción de genes silenciados. Las marcas en gris representan genes cuyo sitio de inicio de la transcripción (TSS) está en la transición entre dos tipos de cromatina (Modificado de Filion et al., 2010).

Un año después, Kharchenko hizo un análisis similar, estudiando 18 modificaciones de histonas, y, en función de ellas, estableció 9 patrones combinatorios. Los análisis integrativos con otros datos, como los de proteínas de la cromatina no histonas o hipersensibilidad a DNAsa I entre otros, revelaban patrones comunes de cromosomas, genes, elementos regulatorios y otros dominios funcionales (Kharchenko et al., 2011).

El estado 1 (Figura 6, rojo) recoge los promotores activos y los TSS, marcados por un enriquecimiento en H3K4me3/me2 y H3K9ac.

El estado 2 (Figura 6, morado), contiene preferentemente la marca de elongación transcripcional H3K36me3, y se encuentra sobre todo en regiones exónicas de genes transcritos.

El estado 3 (Figura 6, marrón), típico de intrones, se distingue por estar enriquecido en H3K27ac, H3K4me1 y H3K18ac.

El estado 4 (Figura 6, coral) contiene H3K36me1, pero se caracteriza principalmente por la carencia de H3K27ac.

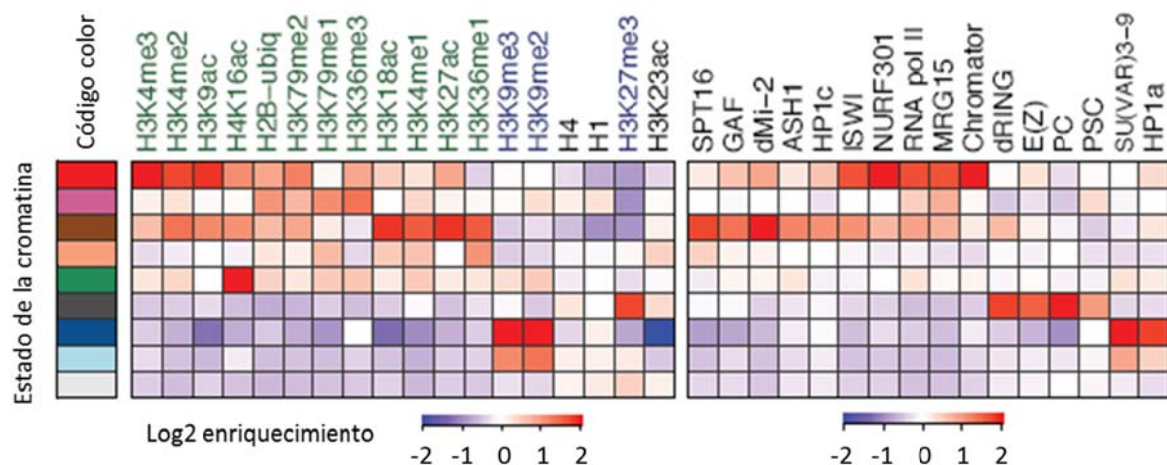
El estado 5 (Figura 6, verde) está enriquecido en genes del cromosoma X, como se observa por la alta concentración de H4K16ac en combinación con H3K36me3 y otras marcas propias del estado 2 (un patrón combinatorio asociado al *dosage compensation* en células de machos (Larschan et al., 2007)).

El estado 6 (Figura 6, gris oscuro) se corresponde con los sitios donde la represión se realiza por el grupo Polycomb.

El estado 7 (Figura 6, azul oscuro) recoge los dominios pericentroméricos de la heterocromatina, así como el cromosoma 4, enriquecidos en H3K9me2/me3.

El estado 8 (Figura 6, azul claro) recoge otra región heterocromática que contiene niveles moderados de H3K9me2/me3, que ocupan dominios de brazos eucromáticos en autosomas, así como en el cromosoma X en ambos sexos.

Por último, el estado 9 (Figura 6, gris claro) contiene otras zonas silenciadas.



**Figura 6. La cromatina se divide en función de 9 tipos de estados.** Representación de los estados de la cromatina (filas) definidos por unos patrones combinatorios de enriquecimiento (rojo) o depleción (azul) de marcas de la cromatina y proteínas cromosomales (Modificado de Kharchenko et al., 2011).

## **Modificación posttranscripcional de histonas**

De estos estudios se deriva que uno de los principales mecanismos por los que la estructura y la función de la cromatina están reguladas es a través de las modificaciones posttranscripcionales de histonas (PTM). La cola N-terminal de las histonas (aunque principalmente lo hacen aquí, también hay sitios en la C-terminal y en sus dominios centrales) sufre modificaciones covalentes, como acetilaciones, metilaciones, fosforilaciones, ubiquitinaciones, ADP-ribosilaciones... Estas modificaciones ejercen su acción por dos mecanismos principales, afectando directamente a la estructura global de la cromatina, tanto a corta como a larga distancia, así como regulando, tanto positiva como negativamente, la unión al DNA de otras moléculas efectoras.

## **Introducción**

A pesar de la gran variedad de modificaciones que han sido descritas, las más estudiadas respecto a su influencia en la transcripción génica han sido la acetilación, metilación y fosforilación.

La **acetilación** se da en lisinas de todas las histonas, neutralizando la carga positiva de este residuo. Tiene un papel fundamental en la regulación transcripcional (Struhl, 1998) ya que se asocia con cromatina descondensada y con accesibilidad al DNA (Shahbazian and Grunstein, 2007), mientras que la hipoacetilación conlleva condensación y represión transcripcional. La reacción la catalizan las acetiltransferasas, frecuentemente presentes en activadores transcripcionales, y puede ser revertida por las desacetilasas de histonas. De la misma manera, muchos represores transcripcionales contienen actividad desacetilasa (Strahl and Allis, 2000). Los *enhancers* y sobre todo los promotores de los genes están enriquecidos en esta marca, facilitando el acceso a factores de transcripción (Wang et al., 2008).

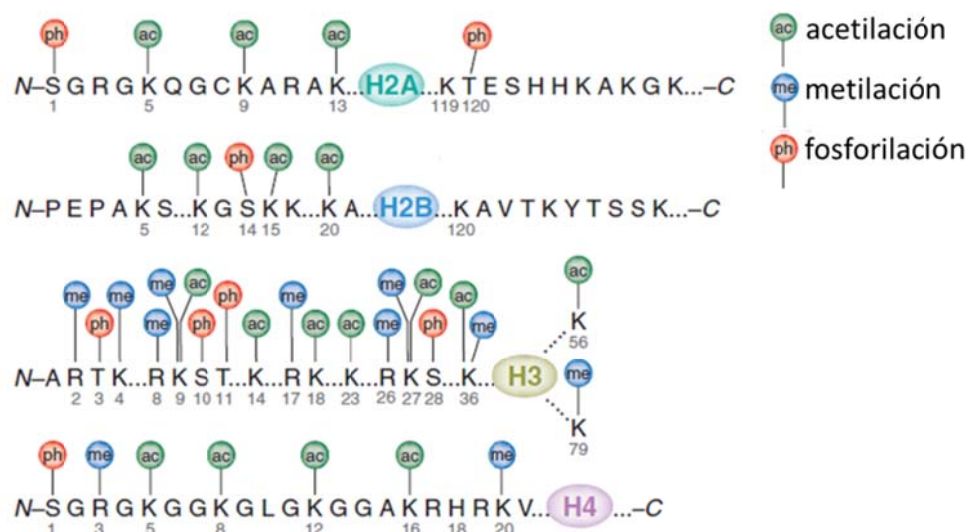
La **metilación** se produce principalmente sobre lisinas, que pueden ser mono, di o trimetiladas, o argininas, con mono o dimetilación, simétrica o asimétrica, y se da sobre todo en H3 y H4, sin modificar la carga. Es llevada a cabo por metilasas, y eliminada por desmetilasas. Inicialmente fue asociada con el silenciamiento de genes, como es el caso de las marcas H3K27me3 asociada al grupo Polycomb o H3K9me2/3 con HP1. No obstante, también se ha estudiado esta PTM en relación con la activación transcripcional, tal como H3K4me3 realizada por el complejo trithorax (trxG) que marca transcripción activa o H3K36me3 que está implicada en la elongación de la transcripción y está presente en toda la región transcrita (Barski et al., 2007, Wang et al., 2008).

La **fosforilación** ocurre en serinas, treoninas y tirosinas de manera preferente aunque no exclusiva en la cola N-terminal, aportando cargas negativas a la histona. Las proteínas kinasas y fosfatasas son las que se encargan de añadir y eliminar la modificación, respectivamente. Inicialmente se relacionó la fosforilación de las histonas con la condensación del DNA durante la mitosis, tanto en H1, como en H3 (Bradbury, 1992, Koshland and Strunnikov, 1996). Uno de los ejemplos mejor estudiados es la fosforilación de la serina 10 de la histona 3 (H3S10ph), que se produce durante la mitosis en muchos eucariotas y se correlaciona con la condensación y segregación cromosómica en mitosis y meiosis (Hendzel et al., 1997). La familia de las Aurora kinasas, muy conservadas en diferentes organismos, son las responsables de esta fosforilación durante estos procesos.

A pesar del evidente papel que juega en la fase divisoria del ciclo celular también se han caracterizado ejemplos en los que la fosforilación de histonas puede estar implicada en la remodelación de la cromatina y la expresión génica. Se vio que la estimulación con factores mitogénicos que provocan la transcripción de genes de respuesta rápida, como *c-fos* y *c-jun*, producía la fosforilación de H3 (Mahadevan et al., 1991), y unos años después se concretó que se daba en el residuo S10 en la cromatina interfásica. Esta modificación no es realizada por Aurora kinasas sino por Jil-1/MSK1-2 (*Drosophila* y mamíferos, respectivamente). El mejor ejemplo que describe la participación de Jil-1 y H3S10ph en la expresión génica viene del estudio de la doble expresión que experimentan los genes del cromosoma X en machos de *Drosophila*, donde se encuentra claramente enriquecida esta proteína y la marca (Regnard et al., 2011).

Esto generó la cuestión de cómo una misma modificación puede desencadenar procesos tan dispares como la condensación de la cromatina en mitosis, y su descompactación en interfase. Surgió la hipótesis de que la marca pudiera actuar combinada con otras. De hecho, se vio que en su efecto mitótico, actuaba junto con H3S28ph (Goto et al., 2002), modificación que también es llevada a cabo por las Aurora kinasas. De igual manera, en el estudio del papel de H3S10ph en interfase, se vio que la presencia de H4K16ac en genes del cromosoma X en machos era lo que provocaba que Jil-1 se uniera el doble en ellos que en otras regiones transcripcionalmente activas y, por tanto, mediara en la activación de estos genes. Por tanto, la combinación de H3S10ph y H4K16ac sería lo que desencadenaría los efectos en la transcripción génica (Regnard et al., 2011).

Así, el conjunto de las modificaciones de histonas y su correlación con ciertos procesos y estados génicos ha llevado a algunos autores a plantearse que existe un “Código de Histonas” (Figura 7), en el que los cambios actúan por separado, secuencialmente o en combinación para dar un lenguaje o código que se lee por proteínas especializadas para facilitar diferentes eventos que se desencadenan gracias a ellas (Strahl and Allis, 2000).



**Figura 7. Modificaciones covalentes de las histonas.** Representación de las modificaciones posttranscripcionales más estudiadas en función de los residuos hallados en los aminoácidos de las histonas. En verde acetilaciones; en rojo fosforilaciones; en azul metilaciones (Modificado de Bhaumik et al., 2007).

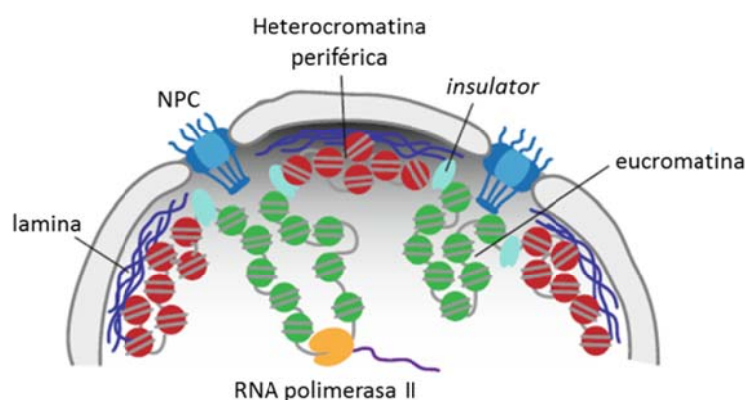
### **Interacción de la cromatina con la lamina**

Así como los tipos de cromatina correlacionan con PTMs, la localización respecto a estructuras nucleares también se tiene en cuenta para la clasificación. En esta categorización del genoma se puede distinguir, por ejemplo, la unión a la lamina (en el caso de la cromatina negra en la primera clasificación, o del estado 9, en la segunda).

## Introducción

Se sabe que la lamina nuclear interacciona con el DNA y con histonas (Taniura et al., 1995, Marshall, 2002), así como que la proteína LAM de la lamina regula la expresión de genes (Pickersgill et al., 2006). Diversos estudios abalan el hecho de que la cromatina periférica está silenciada, y que se localiza de manera más interna cuando se va a transcribir. Por ejemplo, cuando el gen *cystic fibrosis transmembrane conductance receptor (CFTR)* no se transcribe, está asociado a la periferia nuclear. Sin embargo, cuando se expresa este gen se mueve hacia el interior del núcleo y se asocia con la eucromatina interna (Zink et al., 2004). Lo mismo ocurre con los genes implicados en la biogénesis de lípidos, que al sobreexpresarse durante la adipogénesis se mueven de la periferia hacia el interior (Szczerbal et al., 2009). Además, algunos estudios confirman que la localización ectópica de genes en la lamina nuclear reduce la expresión de los mismos, y que su posterior colocación en el interior restaura la actividad perdida (Reddy et al., 2008). No obstante, mediante el posicionamiento artificial de determinados cromosomas en la periferia nuclear en células humanas se observó que la expresión de algunos genes se suprimía, pero que esto no era un hecho general, ya que otros genes no experimentaban este silenciamiento (Finlan et al., 2008). Por tanto, a pesar de la evidencia de que los genes reprimidos están en la periferia, este experimento indicaba que esta localización no es incompatible con la transcripción.

Por otro lado, en levaduras, la activación de algunos genes, como *Gal*, provocaba su relocalización del interior nuclear a la periferia, concretamente a los complejos de los poros nucleares (NPC) (Casolari et al., 2004). Este hecho indicaba que en la periferia nuclear, al menos en levadura, podían existir ambos tipos de regiones, transcripcionalmente activa e inactiva. Quizás la separación de ambas se produzca por proteínas *insulators*, que se sabe que se localizan en estos NPC (Ishii et al., 2002), impidiendo la propagación del silenciamiento de la heterocromatina a la eucromatina (Figura 8). En *Drosophila* también hay evidencias de que la periferia puede favorecer la expresión de ciertos genes, como los del cromosoma X en machos para duplicar su nivel de expresión (Akhtar and Gasser, 2007). Se vio que MSL, una de las proteínas implicadas en el *dosage compensation*, interaccionaba con componentes de los NPC, y que en ausencia de éstos, los factores involucrados en el DCC no se asociaban con los genes del cromosoma X (Mendjan et al., 2006).



**Figura 8. Modelo de organización de la cromatina en la periferia nuclear.** Representación de la envoltura nuclear, que es una doble membrana perforada por complejos de los poros nucleares (NPC), a los que está unida la eucromatina. La membrana interna, lamina, está compuesta de proteínas filamentosas, que unen la heterocromatina periférica. Las proteínas *insulators* separan la eucromatina de la heterocromatina.

El conocimiento de cómo se organiza la cromatina dentro del núcleo ha experimentado un notable avance. Se ha pasado de diferenciar simplemente zonas transcripcionalmente activas y silenciadas a poder detallar diferentes tipos de cromatina en función de las características que presentan en común, así como las interacciones que ocurren entre zonas separadas en el genoma dividiéndolo en dominios.

De todo lo expuesto hasta ahora se deriva la importancia que tiene la estructura de la cromatina así como todos los elementos que intervienen en esta organización a la hora de regular la expresión de genes. Por ello, pasaremos a detallar este proceso, qué elementos los llevan a cabo y qué mecanismos están implicados en la regulación del mismo.

## **1.2. Regulación de la expresión génica**

### **1.2.1. Elementos básicos de la transcripción**

Cada célula establece unos patrones individuales de expresión de los genes durante la diferenciación y desarrollo. Por tanto, es necesario un control espacio-temporal de la transcripción de estos genes. La transcripción es el proceso en el que se produce RNA a partir de una cadena molde de DNA. La enzima que cataliza esta reacción es la RNA polimerasa. En las células eucariotas se distinguen tres tipos, y la RNAPolIII es la responsable de transcribir los mRNA, que son los que servirán de molde para dar lugar a proteínas. Durante la transcripción, forma un complejo de pre-iniciación (PIC) multiproteico con otros factores denominados factores generales de la transcripción (GTF), ya que los necesita para reconocer el promotor basal.

El promotor basal es la mínima secuencia de DNA suficiente para dirigir el inicio de la transcripción por la maquinaria de la RNAPolIII, mientras que se denomina promotor proximal a la zona que cubre unos 250 nucleótidos a un lado y a otro del sitio de inicio de la transcripción (Butler and Kadonaga, 2002). El promotor basal típicamente incluye el TSS y se extiende 35 nucleótidos más allá, tanto *upstream* como *downstream*. Contiene unas secuencias específicas (Butler and Kadonaga, 2002) a través de las que interactúa con factores externos para regular la transcripción génica. Se han visto diferentes tipos de motivos, como la caja TATA, el elemento *Initiator*, el TFIIB *recognition element* o el *downstream core promoter element*. Cada uno de ellos tiene su función específica, y ninguno se encuentra en todos los promotores, por lo que no hay ninguno que sea universal. La combinación de estos motivos y su reconocimiento consiguen modular los niveles de expresión (Kadonaga, 2012).

La unión del PIC al promotor basal provoca la separación de las cadenas de DNA para poder comenzar a transcribir, fase conocida como iniciación. La siguiente fase es la de elongación, que



## **Introducción**

consiste en el avance de la maquinaria transcripcional a lo largo del gen. A medida que avanza se va sintetizando la molécula de RNA mensajero. Finalmente, la RNAPolIII llega al extremo 3' del gen y se separa del DNA y del RNA, constituyendo la fase de terminación.

### **1.2.2. Mecanismos implicados en la regulación de la transcripción génica**

La determinación del estado transcripcional se modula a través de una combinación de mecanismos, entre los cuales no solo cabe destacar las diversas modificaciones que pueden sufrir las histonas como se ha explicado en el apartado 1.1.1.5., sino también la participación de los *enhancers*, el silenciamiento llevado a cabo por elementos de respuesta a Polycomb (PRE) y el aislamiento de regiones reguladoras por proteínas *insulators*.

#### **1.2.2.1. Enhancers**

Los primeros *enhancers* descritos eran unas secuencias víricas repetitivas del SV40, capaces de activar la expresión de un constructo con el gen de la  $\beta$ -globina de conejo insertado en un cultivo de células HeLa (Banerji et al., 1981). Esta activación era independiente de la orientación y localización del *enhancer* dentro del plásmido. Después se han caracterizado en diferentes organismos (Arnold et al., 2014, Villar et al., 2015), considerándose como piezas clave en la regulación de la expresión génica en todos los eucariotas.

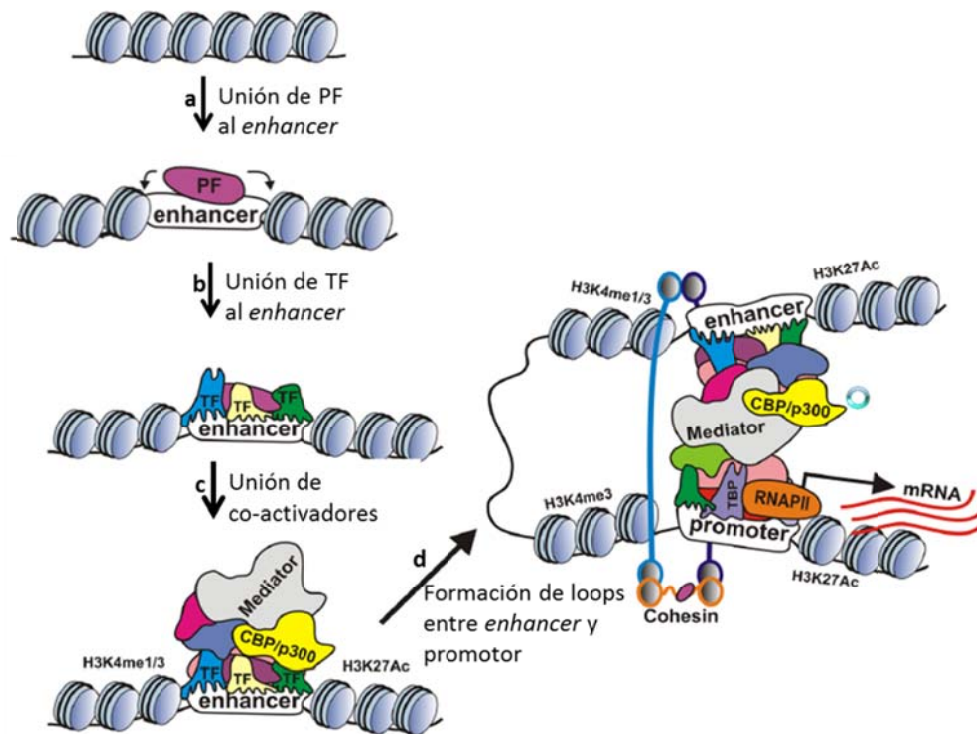
De los estudios de SV40 se extrajeron las características comunes (Banerji et al., 1981, Moreau et al., 1981, Treisman and Maniatis, 1985). Se pueden definir como secuencias no estereotipadas de DNA que son capaces de regular la transcripción de un gen a pesar de estar a miles de bases de distancia, tanto *upstream* como *downstream*, y en cualquier orientación. Pueden estar localizados en regiones intra o intergénicas, en intrones e incluso en exones de los genes (Erokhin et al., 2015).

A pesar de que múltiples *enhancers* pueden interactuar con múltiples promotores, se ha detectado una especificidad entre ciertos *enhancers* hacia ciertos promotores. Uno de los primeros estudios sobre esta compatibilidad se realizó mediante la expresión de dos genes vecinos, *gooseberry* (*gsb*) y *gooseberryneuro* (*gsbn*). Se observó que aunque dos *enhancers* (GsbE y GsbnE) estaban localizados entre los TSS de ambos genes (y, por tanto, una activación cruzada pudiera ocurrir), GsbE solo podía activar al promotor de *gsb*, y GsbnE solo al de *gsbn* (Li and Noll, 1994). La preferencia de un *enhancer* por un promotor en concreto parece depender de varios factores, como la compatibilidad bioquímica, la distancia entre ellos, los motivos que presenta el promotor, la arquitectura 3D de los cromosomas y del estado de la cromatina local, así como de elementos externos, tal como los *insulators* (van Arensbergen et al., 2014).

Se han propuesto varios modelos sobre cómo se puede hacer el contacto entre un *enhancer* y un promotor cuando están físicamente muy alejados. Inicialmente se habló del modelo de escaneo, en el que un *enhancer* une la RNAPolIII y GTFs, desencadenando el avance de la polimerasa en ambos sentidos hasta encontrarse con el promotor, creando el PIC (Moreau et al., 1981), o bien mediante la propagación de estructuras de la cromatina que extenderían un estado abierto de la misma hasta el promotor. Sin embargo, pronto se rechazó esta hipótesis para buscar otras en las que hubiera contacto físico entre ambos elementos mediante proteínas ya que éste había sido confirmado en varios estudios. El modelo que más evidencia ha recaudado es el de la formación de *loops* de cromatina entre ellos cuando interaccionan, que dejarían fuera toda la secuencia que queda en medio (Kagey et al., 2010, Mifsud et al., 2015). Los avances en el campo gracias a las técnicas 3C han permitido afirmar que los *enhancers* están en proximidad física con el gen que están activando. Como previamente hemos señalado, uno de los estudios pioneros es el del gen  *$\beta$ -globina*. A pesar de que el LCR se sitúa a cierta distancia del propio gen, se ha visto que colocalizan en células donde se está expresando, pero no en aquellas en las que está inactivo (Tolhuis et al., 2002). La conclusión de éste y otros estudios refuerzan la teoría de que la interacción *enhancer*-promotor se realiza a través de estos bucles (Rao et al., 2014).

Con el fin de comenzar la transcripción, el *enhancer* recluta un complejo proteico iniciador que facilita la accesibilidad a su DNA desplazando los nucleosomas, mediante la aparición de PTMs activadoras y abriendo la cromatina localmente. De esta manera, los factores de transcripción pueden unirse, permitiendo así el reclutamiento de co-activadores que no pueden establecer contactos directamente con el DNA, dando un complejo activo. La unión de estos últimos es lo que consiente la interacción entre *enhancer* y los GTF de promotores (Figura 9) (Palstra and Grosveld, 2012, Erokhin et al., 2015).

El cofactor transcripcional Mediator, acompañado del complejo de las cohesinas, tienen un papel directo en la comunicación entre *enhancer* y promotor (Figura 9). El primero establece contacto físico entre factores de transcripción de ambos elementos, transmitiendo la señal de activación desde el *enhancer* al promotor. Esta interacción promueve la formación de un *loop* de cromatina, y el complejo de las cohesinas lo estabiliza formando una estructura en forma de anillo entre los elementos que interaccionan (Kagey et al., 2010).



**Figura 9. Activación de la transcripción génica.** Esquema representativo del contacto entre *enhancer* y promotor. **a/** Primero se reclutan factores de iniciación (*Pioneer Factors*, PF) en el *enhancer* facilitando la accesibilidad a su DNA. **b/** Tras ello, se unen los factores de transcripción (TF) **c/** Unión de factores co-activadores que se unen a los TF previamente unidos para formar un complejo activo. **d/** Contactos entre los TF del *enhancer* activo y los GTF del promotor, mediante *loops* gracias a Mediator, y las cohesinas los refuerzan formando la estructura de anillo alrededor del DNA (Modificado de Erokhin 2015).

Sin embargo, no solo Mediator y el complejo de las cohesinas participan en estos contactos, también se ha visto que CTCF media en el proceso. Esto sugiere que la activación génica también puede ser regulada por la proteína *insulator* CTCF (Sekimata et al., 2009, Ong and Corces, 2014). A pesar de su conocida actividad *enhancer-blocking* cuando se encuentra en la frontera entre dominios, su presencia dentro de los mismos puede favorecer la activación génica. Varias investigaciones corroboran esta función de CTCF. Por ejemplo, en uno de estos estudios se analizaron las interacciones entre promotores y *enhancers*, y se vio que casi el 80% de las mismas no se impedían al tener uno o varios sitios de unión de CTCF entre ellos y muchos de los *enhancers* estaban enriquecidos en sitios de unión de CTCF lo que indicaba que la proteína *insulator* estaba implicada en regulación transcripcional (Sanyal et al., 2012). Esta dualidad de funciones de CTCF, opuestas entre sí, se explica porque las proteínas *insulators* están relacionadas con la regulación de la expresión génica en función del contexto genómico, y no con la activación o silenciamiento génico de manera exclusiva. Como en el caso de Mediator, el complejo de las cohesinas estabiliza los contactos realizados por CTCF a través de los anillos que establecen.

La cuestión de cómo el *enhancer* activa la transcripción una vez se ha puesto en contacto con el promotor no está del todo clara. Un mecanismo sugerido sería regulando la transición de la RNAPolIII pausada a la activa en el promotor proximal, ya que se ha visto que la ausencia de LCR del gen  $\beta$ -globina disminuía la elongación mientras que se mantenía la unión del PIC y la RNAPolIII en el promotor (Sawado et al., 2003). Por otro lado, también se ha sugerido que sería el *enhancer* el encargado de reclutar la RNAPolIII y los GTF y entregarlos al promotor. Este mecanismo se ha demostrado *in vitro* pero no *in vivo*. En cualquier caso, se cree que el *enhancer* puede tener una labor secuencial que englobe diferentes mecanismos, desde mantener el estado abierto, reclutar factores de transcripción específicos y ensamblar el PIC en el promotor del gen mediante *loops* de cromatina, así como liberar la RNAPolIII pausada reclutando factores de elongación (Palstra and Grosveld, 2012).

### 1.2.2.2. Elementos de respuesta a Polycomb

Como hemos señalado en el apartado anterior, la comunicación entre *enhancer* y promotor, gracias a factores de transcripción y RNAPolIII, establece el patrón inicial de la expresión génica. En algunos casos, cuando los TFs dejan de estar presentes, estos estados se tienen que mantener mediante sistemas epigenéticos de memoria celular, en los que las proteínas del grupo Polycomb y trithorax tienen un papel destacado.

En 1947 Lewis descubrió el gen *Polycomb* (*Pc*) en *Drosophila*, un regulador del desarrollo requerido para mantener la represión de los genes homeóticos del complejo Bithorax (BX-C). Con el paso del tiempo se ha visto que esta proteína, así como otras que forman parte de la misma familia, no solo juegan un papel en el silenciamiento de genes Hox y en el desarrollo (Lewis, 1978), sino que también intervienen en otros procesos como el ciclo celular (Oktaba et al., 2008), diferenciación (Pasini et al., 2007), respuesta al estrés (Landais et al., 2014) o como sistema de memoria, estableciendo, manteniendo y transmitiendo el silenciamiento epigenético de la cromatina (Berry et al., 2015).

Estas proteínas forman complejos conservados a lo largo de la evolución, sobre todo los dominios con función clave (Di Croce and Helin, 2013). Mientras que en metazoos en general podemos diferenciar 2 complejos represivos de Polycomb (PRC), PRC1 y PRC2, en *Drosophila* hay 3 (Geisler and Paro, 2015).

El PRC2 está compuesto por cuatro proteínas: *Enhancer of Zeste* (*E(z)*), *Extra sex combs* (*Esc*), *Supresor of Zester 12* (*Su(z)12*) y *p55* (*Nurf55* o *Caf1*). *E(z)* es el componente catalítico, que trimetila H3K27 en nucleosomas adyacentes creando H3K27me3 (Cao and Zhang, 2004), marca característica de los genes regulados por PcG y que es reconocida por la proteína *Pc*.

El PRC1 contiene *Polycomb*, *Polyhomeotic* (*Ph*), *Posterior sex combs* (*Psc*) y *Sex combs extra o dRing1* (*Sce/dRing1*). *Pc* tiene un cromodominio que reconoce la marca H3K27me3, anclando el

## **Introducción**

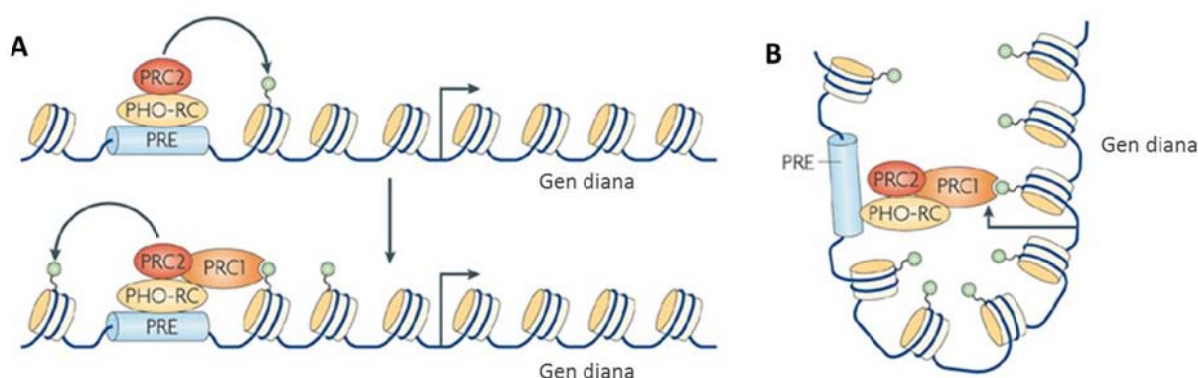
complejo a la cromatina. Sce produce H2AK118ub1, que impide la elongación de RNAPolIII, e incluso parece también estar implicada en el reclutamiento de miembros de PRC2. Además, los eventos mediados por PRC1 compactan la cromatina para limitar el acceso de factores activadores.

El tercer complejo presente en *Drosophila* es *Pho-repressive complex* (PHO-RC) formado por *Pleiohomeotic* (Pho), que es la única de todas estas proteínas con capacidad de unirse directamente al DNA, y dSfmbt.

Los genes Hox también pueden ser diana de otro grupo proteico con acción antagónica a la de PcG, el grupo trithorax. Este grupo es requerido para el mantenimiento del estado activo de un gen, tanto si es diana de PcG como si no lo es (Kingston and Tamkun, 2014). Lo que determina el estado transcripcional del gen que ambos pueden regular son las PTMs que estos grupos producen sobre él. Desde los primeros estudios se sugirió que las modificaciones realizadas por trxG pueden inhibir la actividad de PRC2, y, por tanto, impiden la represión (Klymenko and Muller, 2004, Papp and Muller, 2006). Diversas investigaciones constatan este hecho. En una de ellas se observó que la di o trimetilación de H3K36, realizada por una de las proteínas de trxG, Ash1, no colocalizaba con H3K27me3 en la misma histona, e incluso impedía la acción de PRC2 (Yuan et al., 2011). Por otra parte se describió que H3K4me1/me2/me3 producida por otra de las proteínas del grupo, trx, imposibilitaba la unión de p55 en los sitios donde se encontraba (Schmitges et al., 2011). Según esto, la presencia de estas PTMs impide la actividad de PRC2, y por tanto, la posterior adhesión de PRC1. La inhibición de la función de PRC2 mediante estas marcas es un hecho general entre diferentes especies, como humanos, ratón o *Drosophila*, lo que sugiere que es un mecanismo de regulación génica altamente conservado (Geisler and Paro, 2015). Sin embargo, análisis de ChIP-chip de genes de los complejos Antennapedia (ANT-C) y Bithorax confirmaron que estos genes contenían tanto proteínas de PcG como de trxG simultáneamente (Beisel et al., 2007). En mamíferos existen dominios bivalentes, con promotores que contienen tanto H3K4me3 como H3K27me3 (Voigt et al., 2013). Inicialmente se asoció esta doble marca con la pausa de genes implicados en el desarrollo que necesitan una activación rápida, aunque la teoría actual es que H3K27me3 se establece tras el depósito de H3K4me3, para impedir que otras metiltransferasas activen estas regiones (Denissov et al., 2014).

En el genoma, los complejos de proteínas PcG son reclutados a *Polycomb Group Response Elements* (Figura 10), que son secuencias de DNA que pueden tener hasta unos pocos cientos de pares de bases. Cabe señalar que los PRE no son siempre diferenciables de otros elementos del genoma, los *Trithorax Response Elements* (TRE). Estos son los elementos que unen las proteínas trxG. Ya hemos señalado que las proteínas que forman estos complejos pueden colocalizar, de manera que PREs y TREs se solapan, al menos parcialmente (Tillib et al., 1999), tal como revelan las inmunoprecipitaciones de cromatina (Ringrose and Paro, 2007). Contienen algunos motivos de unión para proteínas como GAGA, Pho, Pho-like, zeste o *dorsal switch protein 1* (DSP1) (Muller and Kassis, 2006). De estos, Pho y Pho-like se han caracterizado como componentes de PcG, y son los que establecen la unión con el DNA para posteriormente reclutar en algunos sitios al resto de proteínas del grupo que están participando en el silenciamiento (Figura 10). Los PREs se pueden encontrar a

varias Kb de distancia del promotor del gen que silencian (Ringrose and Paro, 2007) y su efecto silenciador no depende de si está *upstream* o *downstream* del gen ni de su orientación.



**Figura 10. Polycomb response element.** Representación del silenciamiento génico por proteínas del grupo de Polycomb. **A/** Inicialmente, PHO-RC reconoce PREs y recluta en ellos a PRC2, que metila H3K27 en los nucleosomas adyacentes. La acumulación local de H3K27me3 atrae, une y estabiliza PRC1. Su reclutamiento a PRE puede requerir contacto directo con PHO-RC y/o PRC2. **B/** PRC1 unida a PRE interacciona con H3K27me3 para estabilizar el contacto con el gen. La formación del *loop* une PcG con la región transcrita, donde PRC2 puede extender H3K27me3 ampliando así el silenciamiento de la zona (Modificado de Simon and Kingston, 2009).

Una de las primeras evidencias que relacionaban a los PREs con la regulación transcripcional y la organización nuclear se obtuvo en *Drosophila*. La mayoría de genes del complejo Bithorax que unían proteínas PcG interaccionaban físicamente mediante contactos a larga distancia cuando estaban silenciados. Se ha visto mediante estudios 4C que estos contactos se establecían para estabilizar el mantenimiento del silenciamiento epigenético, y que dependían de la organización en 3D de estos genes (Tolhuis et al., 2011). Tras la reactivación transcripcional artificial de este complejo de genes, se detectó una conformación diferente, en la que los PREs y los promotores activos perdían el contacto (Lanzuolo et al., 2007, Sexton et al., 2012). Las proteínas PcG interaccionan entre ellas para facilitar el silenciamiento de genes y tienden a agregarse en lugares concretos en vez de estar dispersas por el núcleo, formando *Polycomb bodies* (*Pc bodies*). De esta manera, actúan como estructuras en las que las dianas de PcG se agrupan y son correprimidas, a pesar de estar separadas en la secuencia lineal de DNA. Esta idea fue confirmada mediante un estudio de los genes *Antennapedia* (*Antp*) y *Abdominal-B* (*Abd-B*) en *Drosophila*. Se observó que estos genes, dianas de PcG, a pesar de estar localizados a 10Mb uno de otro, colocalizaban cuando estaban inactivos y la generación de una mutación en uno de ellos desencadenaba un debilitamiento del silenciamiento del otro (Bantignies et al., 2011). Se había señalado que los PREs interactuaban entre ellos para mediar la configuración que permitiera el silenciamiento génico. Sin embargo, un estudio demuestra que los PREs no son los responsables de los contactos entre genes silenciados, sino que se establecen a través de elementos *insulators* situados a su lado (Li et al., 2011).

## Introducción

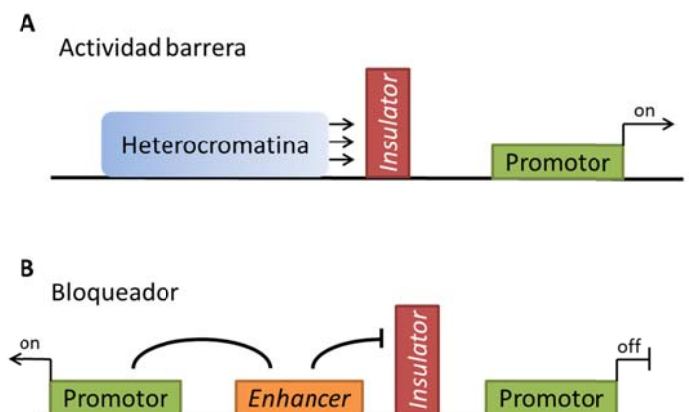
### 1.2.2.3. Insulators

La primera vez que se sugirió la posibilidad de elementos con función *insulator* fue en 1932. Sin embargo, hasta 1985 no se examinaron en detalle, y se vio que había unas secuencias de DNA particulares que delimitaban y aislaban una región de la cromatina en el locus *heat shock* (Udvardy et al., 1985). Desde entonces se han detectado muchas otras secuencias, y a pesar de que fue descubierto en *Drosophila*, después se han ido caracterizando en otros organismos como levaduras y vertebrados (Li and Stamatoyannopoulos, 1994, Donze et al., 1999).

Estas secuencias fueron descritas como elementos que prevenían la comunicación indebida entre regiones genómicas, y que delimitaban cromatina activa y silenciada. Así, se han definido clásicamente por presentar una de las dos siguientes propiedades (Gerasimova and Corces, 2001):

-Actividad barrera: se sitúan en los bordes que separan la cromatina activa de la silenciada, previniendo el silenciamiento de genes activos por el bloqueo de la expansión de la heterocromatina (Figura 11A).

-Bloqueador: impiden la comunicación entre elementos regulatorios de la expresión de genes cuando están situadas entre ambos, como en el caso de *enhancer* y promotores. Se dice que lo hacen de manera direccional ya que no silencian o desactivan el *enhancer*, pues todavía puede activar un gen del que no esté separado por un *insulator*, contribuyendo de esta manera a la correcta expresión temporal y espacial de genes (Figura 11B).



**Figura 11. Funciones clásicas de *insulators*.** **A**/Representación de la actividad de los *insulators* como barrera para la expansión de la heterocromatina (actividad barrera). **B**/Representación de la función de los *insulators* como interruptores de la comunicación entre un *enhancer* y un promotor cuando se sitúan entre ambos (bloqueador) (Modificado de Heger and Wiehe, 2014).

Estas secuencias que acabamos de describir no tienen efecto por sí solas, sino que necesitan de la unión de proteínas *insulators*, ya que esta combinación es necesaria para su correcta actividad. Cada una de estas proteínas interacciona con una red particular de secuencias *insulators*, que generalmente quedan definidas por el contexto de cromatina y la localización genómica. No obstante se pueden solapar con otras proteínas en estos sitios (Negre et al., 2010, Schwartz et al., 2012).

*Drosophila* es el animal modelo en que más proteínas de este tipo se han descrito: *Suppressor of Hairy wing* (Su(Hw)) (Roseman et al., 1993), *Boundary element-associated factor of 32kDa* (BEAF) (Hart et al., 1997), *GAGA factor* (GAF) (Ohtsuki and Levine, 1998), *Zeste-White 5* (Zw5) (Gaszner et al., 1999) y *CCCTC-binding factor* (CTCF) (Moon et al., 2005). Recientemente se han identificado otras como Elba, formado por 3 proteínas, Elba1, Elba2 y Elba3 (Aoki et al., 2012), pita y *zinc-finger protein interacting with CP190* (ZIPIC) (Maksimenko et al., 2015). La caracterización de dos nuevas proteínas, *Insulator Binding Factor 1 y 2* (Ibf1/Ibf2) (Cuartero et al., 2014) se realizó en parte durante el transcurso de esta tesis y el trabajo que yo realicé se explica en el apartado 2.1. de resultados.

Sin embargo, solo se conoce una proteína *insulator* en vertebrados, CTCF (Recillas-Targa et al., 2002), aunque ortólogos de proteínas *insulators* de *Drosophila* como GAF se han descrito (Lunyak et al., 2007, Matharu et al., 2010). Ninguna de estas proteínas han sido encontradas en levaduras ni en plantas, y solo unos pocos nematodos poseen CTCF (Heger et al., 2009). TFIIIC también es considerada proteína *insulator* (Van Bortle and Corces, 2012, Kirkland et al., 2013), y es probablemente la más conservada entre eucariotas.

Normalmente la unión de las proteínas a las secuencias *insulators* es necesaria pero no suficiente para la actividad en un lugar específico. Varios cofactores son requeridos para hacer el contacto físico, así como para anclarlas a estructuras nucleares. En *Drosophila*, se han caracterizado clásicamente *Modifier of mdg4* (Mod(mdg4)) y *Centrosomal protein of 190kDa* (CP190), y posteriormente cohesinas y condensinas, que en diferentes combinaciones están presentes en todos los tipos de *insulators*.

La definición de *insulator* (entendido como el complejo de la secuencia de DNA junto con las proteínas) ha ido evolucionando con el tiempo, pasando de ser unos elementos bloqueadores con efectos a nivel local a tener relevancia a nivel de organización genómica y cromosomal. La hipótesis inicial de que los *insulators* organizan la cromatina a través de *loops* deriva de la necesidad de proveer un mecanismo que explique cómo sus dos funciones clásicas se pueden desarrollar en el núcleo (Corces, 1995). Se ha dejado de ver a los *insulators* como un obstáculo para la expresión génica, ya que tanto pueden favorecer como dificultar la transcripción, en función del contexto en el que se encuentran. Un estudio con transgenes en *Drosophila* confirmó que la presencia de una secuencia gypsy entre *enhancer* y promotor, o cuando el promotor estaba flanqueado por dos *insulators*, bloqueaba la comunicación entre ellos. En cambio, cuando había dos secuencias entre ambos se formaba un *loop* entre los *insulators* que acercaba el *enhancer* al promotor, facilitando la activación (Muravyova et al., 2001). La regulación génica a través de la formación de bucles fue posteriormente confirmada en mamíferos, donde al eliminar CTCF se producía una reducción de los *loops* en el gen de  *$\beta$ -globina* en ratón (Splinter et al., 2006), mientras que la adición de una copia ectópica en el mismo gen en células humanas inducía la formación de otros alternativos que afectaban a la comunicación entre *enhancer* y promotor del gen (Hou et al., 2008). Se piensa que cohesinas y condensinas ejercen su mecanismo de acción a la hora de participar en los *loops* gracias a su capacidad de formar las estructuras en anillo, estabilizando los *loops* en cis mediados por las proteínas *insulators* (Hadjur et al., 2009, Nativio et al., 2009, Hou et al., 2010).



## **Introducción**

El estudio de la regulación del gen *Abdominal-B* (apartado 2.1.2.) ha permitido profundizar en esta función de los *insulators*. Son capaces de organizar la estructura de la cromatina de toda la región reguladora a través de contactos lejanos entre los elementos y la región promotora de *Abd-B*. Se ha visto que el *insulator* del Fab7 está en contacto con el del Fab8, así como con el promotor del gen cuando éste está silenciado (Cleard et al., 2006). Se ha detectado por ChIP la presencia de proteínas *insulators* como CP190 o CTCF no solo en los elementos frontera, sino también en la región promotora del gen (Moon et al., 2005, Holohan et al., 2007, Mohan et al., 2007). La interacción entre proteínas *insulators* consigue, por un lado, formar *loops* que dejan fuera dominios enteros y, por otro lado, ponen en contacto elementos presentes en el dominio con el promotor del gen, mediando la correcta expresión génica (Muller et al., 1999, Sipos and Gyurkovics, 2005, Kyrchanova et al., 2008).

Los contactos mediados por los *insulators* no solo facilitan o dificultan la comunicación entre *enhancer* y promotor, sino que también regulan el silenciamiento llevado a cabo por proteínas del grupo de Polycomb. Se ha percibido que una sola copia de gypsy puede impedir la expansión de la represión mediada por proteínas PcG, de acuerdo a su función como bloqueador de la comunicación. En cambio, si hay dos secuencias *insulators* entre el PRE y el promotor del gen, éstas interaccionan entre ellas, restaurándose la represión (Mallin et al., 1998, Comet et al., 2006). Esto ha sido confirmado en experimentos de 3C, en los que no solo se vio la interacción entre las dos secuencias, sino también el contacto directo entre PRE y promotor, así como el enriquecimiento de H3K27me3, que no se observa al dejar solo un *insulator* (Comet et al., 2011).

Aunque inicialmente se había propuesto un papel clave de los PRE del BX-C para mediar el silenciamiento génico mediante contactos (Lanzuolo et al., 2007) diversos análisis funcionales y de interacción muestran que los contactos son independientes de ellos, y que se realizan gracias a proteínas *insulators*. En un estudio se observó que el *insulator* Fab7 estaba en contacto con el promotor del gen *Abd-B* cuando estaba silenciado (Cleard et al., 2006), y años más tarde se demostró la labor que desempeñan los *insulators* para mantener el silenciamiento génico. La asociación de dos dianas de proteínas PcG como son los elementos reguladores de Mcp y Fab7 no dependía de PREs. Las interacciones entre copias lejanas de los mismos eran mediadas por los *insulators* situados junto a los PREs en las regiones reguladoras, y no de los propios PREs, como se derivó de ensayos funcionales y de colocalización física a través de inmunotinción, así como de técnicas 3C (Li et al., 2011).

Por la relevancia que tienen para esta tesis, describiremos algunas de las proteínas con actividad *insulator*, como son CTCF, CP190, cohesinas y condensinas.

### **CTCF**

CTCF fue inicialmente descubierta en vertebrados cuando se estudiaban proteínas que se unían a regiones hipersensibles a nucleasas de la secuencia que flanqueaba el gen *c-myc* de pollo. Cuando faltaba esta proteína, la expresión del gen se veía alterada (Lobanenkov et al., 1990).

Posteriormente, ya fue caracterizada como proteína *insulator*, comprobándose su actividad *enhancer-blocking* en algunos *insulators* de vertebrados. Un ejemplo es el locus  $\beta$ -globina de pollo. Tras descubrirse que esta región podía bloquear la actividad del *enhancer* (Chung et al., 1993), se identificó en él la secuencia responsable de esta función. Demostraron que era el sitio de unión para CTCF, y que precisamente la unión de esta proteína a esta secuencia era necesaria para que desempeñara la actividad (Bell et al., 1999). Posteriormente, a través de tecnología 3C, se observó que CTCF ayudaba a la formación de *loops* que abarcaban elementos como el gen  $\beta$ -globina y el LCR (Splinter et al., 2006). Estos sitios de  $\beta$ -globina quedaban reposicionados tras la formación del *loop* de tal manera que se bloqueaba la comunicación entre el *enhancer* y el promotor del gen y la transcripción génica se reprimía. De hecho, ya previamente habíamos señalado la importancia que tiene la formación de bucles por parte de CTCF en mamíferos para organizar la estructura del genoma estableciendo las fronteras entre dominios (apartado 1.1.1.3.) y como regulador de la transcripción mediando en la comunicación entre *enhancer* y promotor (apartado 1.2.2.1.).

Posteriormente se identificó un ortólogo en *Drosophila* y se comprobó tanto *in vivo* como *in vitro* que en este animal modelo también era capaz de unirse a secuencias *insulators*. A través de ensayos *enhancer-blocking* se comprobó que era esencial para el funcionamiento del Fab8 (Moon et al., 2005). De hecho, se ha detectado que junto con CP190 se encuentra unido a seis de los *insulators* que separan los dominios del complejo Bithorax. Su falta provoca fenotipos de transformaciones homeóticas y la alteración de la expresión del gen *Abd-B* como se ha visto por inmunotinción del nervio cordal en larvas. Todo esto recalca la importancia de CTCF para regular los genes homeóticos, la segmentación abdominal y la regulación del desarrollo de las moscas (Mohan et al., 2007).

Mediante *ChIP-seq* en células Kc se observó que los extremos de los dominios silenciados que contienen H3K27me3 se encontraban enriquecidos en CTCF. De acuerdo a su actividad clásica como barrera, su función en estos dominios sería la de mantener la estructura de la cromatina impidiendo la expansión de la heterocromatina (Bartkuhn et al., 2009). Sin embargo, la disminución de la proteína provocaba una depleción de la marca silenciadora dentro de los propios dominios sin producirse un aumento fuera de ellos. Los niveles de H3K27me3 estaban específicamente reducidos en dominios de Polycomb flanqueados por CTCF, indicando que la pérdida de H3K27me3 era un efecto directo de la falta de CTCF, y no una consecuencia general de la alteración de la estructura de la cromatina. Este hecho recalca la importancia que CTCF tiene a la hora de mantener los dominios de Polycomb tal y como sucede en los *iab* del Bithorax (Van Bortle et al., 2012). Por tanto, es capaz de mediar en el mantenimiento del estado transcripcional de un dominio, no solo para evitar el silenciamiento bloqueando la expansión de la heterocromatina, sino también para que los dominios que no se expresan no se vean afectados por los transcripcionalmente activos.

### CP190

A pesar de que CP190 fue descubierta por su papel en los centrosomas (Kellogg and Alberts, 1992), durante la interfase se encuentra dentro del núcleo (Oegema et al., 1995) y se une a

## **Introducción**

cromosomas politénicos en un gran número de sitios (Whitfield et al., 1995). De hecho, su papel como *insulator* ha sido estudiado de manera amplia y detallada.

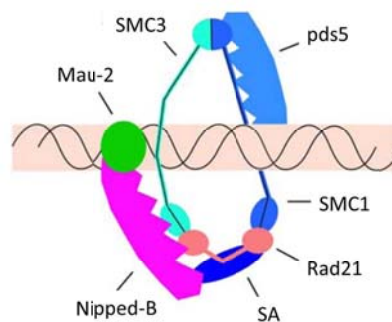
La primera indicación de que CP190 podría estar participando en esta función proviene del estudio del *insulator* gypsy. Se conocía que esta secuencia estaba regulada por Su(Hw) y Mod(mdg4), y un *screening* genético determinó que CP190 potenciaba la acción de esta última. Se vio que CP190 era un constituyente esencial de este *insulator* ya que su presencia era necesaria para que las otras proteínas pudieran bloquear la comunicación entre los elementos reguladores, y que se asociaba físicamente a ambas proteínas, colocalizando con ellas en cromosomas politénicos. No interaccionaba directamente con la secuencia, sino que lo hacía con las proteínas unidas a ella. De hecho, se vio que era necesaria para la formación de los *insulator bodies* en los que estaban Su(Hw) y Mod(mdg4), pues en su ausencia no se agrupaban en estos cuerpos (Pai et al., 2004).

Posteriormente, mediante inmunolocalización en politénicos y coimmunoprecipitación de la cromatina se observó que también interaccionaba con otra proteína *insulator*, CTCF, y se confirmó la presencia de ambas en los *insulators* del complejo Bithorax como hemos señalado (Gerasimova et al., 2007, Mohan et al., 2007). CTCF también se encuentra en los mismos *insulator bodies* en los que están CP190, Su(Hw) y Mod(mdg4). Sin embargo, no interacciona con Su(Hw) ni con Mod(mdg4) como se observa en cromosomas politénicos, que aunque comparten algunos sitios de unión, difieren en su distribución. De hecho, la carencia de estas proteínas no impide que CTCF siga formando parte de los *insulator bodies*; y viceversa, la falta de CTCF no afecta a las otras dos. No obstante, tal como ocurre con las proteínas Su(Hw) y Mod(mdg4), CTCF sí que necesita que CP190 esté presente para formar parte de ellos.

Analizando todo el genoma, se ha observado que se encuentra presente en los sitios de unión de CTCF, BEAF, Su(Hw) y GAF (Negre et al., 2010), y se asocia con la matriz nuclear, promotores activos y fronteras de dominios reprimidos (Ahanger et al., 2013). Además, el hecho de que sea crucial para la formación de *insulator bodies* y hacer que diferentes proteínas que no interaccionan directamente entre ellas se encuentren en los mismos sitios en el núcleo (Pai et al., 2004) hace que se le considere como un cofactor común de los diferentes tipos de *insulators*, con la excepción de Zw5. Además, la importancia de CP190 como responsable de las funciones que desempeñan estas proteínas se comprobó mediante tratamiento con ecdisona, que altera la expresión de algunos genes. Demostraron que el reclutamiento de CP190 a los *insulators* era un mecanismo para regular su actividad, como se vio en el Eip75B. Este *insulator*, a pesar de la presencia de CTCF en él, se mantenía pausado, y solo era activo tras el reclutamiento de CP190. Además, se observó que la unión de CP190 incrementaba la fuerza de las interacciones entre proteínas *insulators* o la frecuencia de las mismas a la hora de formar *loops*, afectando a la regulación de la estructura de la cromatina (Wood et al., 2011).

## **Cohesinas**

El complejo de las cohesinas forma parte de las proteínas para el mantenimiento de la estructura de cromosomas (SMC), y está compuesta por un heterodímero de proteínas SMC (SMC1 y SMC3) y 2 no SMC, Rad21/Scc1 y Stromalin SA/SA2 (nombre que toman las proteínas en *Drosophila* y vertebrados) adquiriendo una forma de anillo (Peters et al., 2008). Son cargadas a los cromosomas por Nipped-B-Mau-2 /Scc2 (NIPBL)-Scc4, atrapándolos en su estructura. Eco y pds5 participan en el mantenimiento de esta unión, hasta que son separadas por wapl y pds5 (Figura 12) (Dorsett, 2009).



**Figura 12. Complejo de las cohesinas.** Representación de las proteínas que forman este complejo (SMC1, SMC3, Rad21 y SA) así como los factores implicados en su carga al DNA (Nipped-B y Mau-2) y una de las proteínas responsables de su mantenimiento y descarga (pds5). Se observa la estructura de anillo que forman atrapando a la molécula de DNA (Modificado de Dorsett et al., 2009).

Este complejo fue inicialmente descrito por su papel en la organización estructural de cromosomas durante la replicación y es esencial para la cohesión de cromátidas hermanas en fase S. En la actualidad ha adquirido relevancia su papel en la organización del genoma y control de la expresión de genes durante interfase (Seitan et al., 2013). Los efectos que las cohesinas tienen en el desarrollo se producen por pequeñas reducciones de las proteínas que no afectan a su función mitótica, proviniendo de la desregulación de la expresión génica, como se ha visto en el caso de algunas cohesinopatías (Ball et al., 2014). Por ejemplo, el síndrome de Cornelia de Lange, caracterizado por dismorfismo facial, anomalías en las extremidades superiores, retraso cognitivo y alteraciones de crecimiento, se produce por una pequeña reducción de NIPBL, aunque también se ha visto en pacientes con mutaciones de *SMC1*, *SMC3* o *Rad21*. Sin embargo, no presentan problemas respecto a la cohesión de las cromátidas hermanas, sugiriendo que las alteraciones del desarrollo se deben al papel que ejercen las cohesinas en la regulación génica.

Varias evidencias apuntan a las cohesinas y proteínas que interactúan con ellas como proteínas *insulators* por las similitudes en las funciones que desempeñan.

Diversos estudios en células de mamíferos confirman la colocalización a lo largo del genoma de CTCF y cohesinas (Parelho et al., 2008, Rubio et al., 2008, Wendt et al., 2008). Muchos de estos sitios

## **Introducción**

son en los que CTCF desempeña su función *enhancer-blocking*, como en el *imprinting control region* (ICR) materno. Basándose en ensayos previos que confirmaban la capacidad de ICR en el locus *H19* de impedir la activación de un gen *reporter* cuando se situaba entre éste y un *enhancer* (Ishihara et al., 2006), se vio que la reducción de la expresión dependía de la presencia de CTCF, pero también de las cohesinas. Esta región contenía sitios de unión para ellas y, sorprendentemente, la falta de cohesinas aumentaba la expresión del gen de igual manera que lo hacía la ausencia de CTCF (Wendt et al., 2008), confirmando que eran necesarias para la actividad *enhancer-blocking* del ICR.

En *Drosophila* no hay colocalización entre CTCF y el complejo de las cohesinas (Holohan et al., 2007) como ocurre en mamíferos. Sin embargo, sí que colocaliza con otra proteína *insulator*, concretamente con CP190 (Bartkuhn et al., 2009), en sitios donde el cofactor no se encuentra con CTCF. Las primeras evidencias de que las cohesinas controlan la expresión de genes y el desarrollo en metazoos provienen de estudios en el gen *cut* de *Drosophila*. Se vio que la mutación de Nipped-B impedía la expresión de este gen, y mediante evidencias genéticas se postuló que esta proteína era crítica para la comunicación a larga distancia entre *enhancer* y promotor (Rollins et al., 1999). Sin embargo, en contra de lo que se esperaba por la relación entre ésta y las cohesinas, se vio que la reducción de distintas subunidades del complejo como Rad21, SA o SMC1 aumentaba la expresión del gen (Rollins et al., 2004, Dorsett et al., 2005).

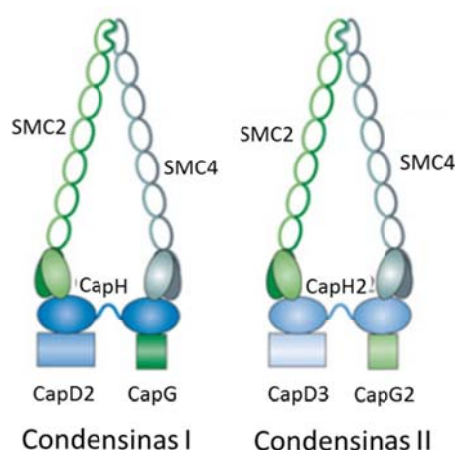
Mediante análisis de ChIP-ChIP se determinó que Nipped-B y cohesinas se situaban preferentemente en sitios activos, y que estaban ausentes de los silenciados (Misulovin et al., 2008). El gen *Abd-B* se está expresando en células Sg4. En esta línea celular, estas proteínas unen *Abd-B* y su región reguladora, pero no otros genes del BX-C silenciados por proteínas PcG. Se vio que los fenotipos que causaba la ausencia de Pc podían ser revertidos mediante la falta tanto de Nipped-B como de Rad21, indicando que ambas proteínas promovían la expresión del gen. Por tanto, comparando este resultado con el del gen *cut*, la función de las cohesinas sobre la expresión génica dependería del contexto (Kennison and Tamkun, 1988, Hallson et al., 2008).

Cuando Nipped-B y cohesinas unen el BX-C, ocupan una zona desde el promotor hasta más allá de la región reguladora *downstream* en el extremo 3'. Esta región está flanqueada por un lado por un sitio de unión de CTCF y por el otro extremo por el Fab7. Este hecho, combinado con que la posición de las cohesinas en mamíferos está determinada por CTCF, sugiere la posibilidad de que diversos tipos de *insulators* ayuden a delimitar las regiones de unión de cohesinas, y que éstas contribuyan a la actividad de los mismos, definiendo dominios activos de la cromatina y determinando las fronteras entre éstos y los silenciados por proteínas PcG (Dorsett, 2009).

Teniendo en cuenta estas observaciones, se ha sugerido que las cohesinas juegan un papel en la regulación de la expresión génica controlando la comunicación entre los elementos reguladores y estableciendo dominios de cromatina. Estas funciones las realizarían mediante la formación de *loops* intracromosomales colaborando en las interacciones mediadas por las proteínas *insulators*.

## Condensinas

Otro complejo de SMC similar al de las cohesinas, el de las condensinas, ha sido tradicionalmente conocido por su función en mitosis en la condensación de cromosomas, pero en los últimos años ha adquirido relevancia por su papel en interfase. Hay dos subtipos, las condensinas I y las condensinas II. Están formados por un heterodímero común, SMC2-SMC4, y otras proteínas características, CapG, CapD2 y CapH para las condensinas I y CapG2, CapD3 y CapH2 para las condensinas II (Figura 13). Solamente el segundo grupo está presente en el núcleo durante todo el ciclo celular, y, por tanto, es responsable de su actividad en interfase (Ono et al., 2004, Bauer et al., 2012).



**Figura 13. Complejo de las condensinas I y II.** Representación de las proteínas que forman estos complejos. SMC2 y SMC4 son comunes en ambos, y tienen 3 diferentes CapH/CapH2, CapD2/CapD3 y CapG/CapG2 en condensinas I y condensinas II, respectivamente. Se observa la estructura de anillo que forman (Modificado de Wood et al., 2010).

Aunque solo estamos en el comienzo de la investigación diversos experimentos confirman que regulan la expresión génica así como la organización nuclear. En *C. elegans* la expresión de los genes de los dos cromosomas X en los individuos hermafroditas tiene que ser reducida a la mitad para equipararse a la de los machos, que solo tienen un cromosoma X. Este proceso es llevado a cabo por un complejo similar al de las condensinas, el *Dosage Compensation Complex*. Se observó que mutaciones de componentes del DCC provocaban un aumento de la unión de la RNAPolIII a los cromosomas X, siendo la primera evidencia de que la compensación ocurre a nivel transcripcional. DCC impide la unión de la polimerasa a los promotores, alterando las interacciones a larga distancia entre *enhancers* y promotores.

Por otro lado, DCC tiene una labor importante en la organización de la estructura de la cromatina. El contacto entre proteínas del complejo es esencial para la regulación génica, y la comunicación a larga distancia entre ellas se realiza a través de *loops*. Además, la secuencia de los cromosomas X está dividida en dominios, y se ha relacionado la formación de las fronteras entre TADs con DCC. La imposibilidad de reclutar este complejo a estos cromosomas impide la regulación génica y también la

## **Introducción**

formación de algunas fronteras entre dominios, o al menos se produce un fuerte debilitamiento de las mismas (Jans et al., 2009). Las más fuertes coinciden con los sitios de unión de DCC; de hecho, la eliminación de estos sitios es suficiente para que las fronteras no actúen correctamente (Crane et al., 2015).

De este estudio se desprende que la capacidad de DCC para mediar en las interacciones entre elementos alejados así como en la estructura de la cromatina es análoga a la función de los *insulators* (Wallace and Bosco, 2013).

Cada vez más líneas de evidencia sugieren que las condensinas juegan un papel importante en la regulación de la estructura de la cromatina, y, por tanto, expresión génica en *Drosophila*. Se observó que la condensina codificada por el gen *barren* estaba relacionada con el mantenimiento del estado silenciado de los genes homeóticos. Se vio por ChIP que la proteína se unía a lo largo del complejo Bithorax, en sitios donde también estaban unidas proteínas PcG, y mediante inmunoprecipitación se demostró la interacción con polyhomeotic, una de ellas (Lupo et al., 2001). Sin embargo, no solo intervienen en el silenciamiento de genes de acuerdo a su papel condensatorio en mitosis. Sorprendentemente, en experimentos de *position effect variegation* se observó que pueden tanto favorecer como inhibir la expresión. Postularon que las condensinas afectan a la expresión génica regulando la expansión de la heterocromatina, posiblemente ejerciendo su acción sobre los *insulators*. Esto confirma que las condensinas son necesarias para mantener un patrón adecuado de expresión génica al menos en determinadas regiones, independientemente de si se trata de silenciamiento o activación (Cobbe et al., 2006).

## **2. DESARROLLO DE *DROSOPHILA MELANOGASTER***

*Drosophila* es un insecto holometabolómico, es decir, que pasa por las fases de larva y pupa antes de la fase adulta. El ciclo completo dura unos 10 días, y comienza cuando, una hembra adulta, tras aparearse, deposita los huevos. El que está fecundado, eclosiona a las 22-24 horas (a temperatura de 25°C) para dar lugar al primer estadio larvario. Tras 24 horas, pasa al segundo y tras otro día al tercer estadio larvario. Después de unas 60-72 horas se convierte en pupa. En esta fase realiza la metamorfosis para dar lugar a un individuo adulto.

Para su correcto desarrollo es necesario que la expresión de genes esté controlada de manera rigurosa. Dos ejemplos que ilustran de manera muy clara esto son la regulación de la expresión de los genes Hox así como el control de las transiciones entre estadios larvarios y la pupación.

## **2.1. Regulación de genes homeóticos**

### **2.1.1. Genes homeóticos**

Los genes homeóticos o genes Hox son una familia de reguladores transcripcionales que controlan diferentes programas de desarrollo a lo largo del eje cabeza-cola de los animales. Fueron inicialmente descubiertos en *Drosophila* formando 2 clusters: Antennapedia y Bithorax (Kaufman et al., 1980). Posteriormente se ha visto que están conservados a lo largo de la evolución (Oulion et al., 2011, Khadjeh et al., 2012, Sheikh et al., 2015).

Presentan colinearidad espacial, es decir, que el orden en el que se expresa cada uno a lo largo del eje cabeza-cola del embrión refleja el orden físico de los genes Hox en los clusters. Están involucrados en la identificación morfológica de cada segmento del cuerpo, y de hecho la alteración de su expresión produce transformaciones homeóticas, en las que la morfología de un segmento se transforma en la de otro (Lewis, 1978). Para controlar con precisión dónde y cuándo se transcriben están acompañados de intrincadas regiones reguladoras.

El estado transcripcional de cada uno de los genes correlaciona con la presencia de determinadas modificaciones de histonas. Es el caso de H3K27me3 en los genes que están inactivos, ya que las proteínas PcG están involucradas en su silenciamiento (Soshnikova and Duboule, 2009). De hecho, se ha visto que todos los genes Hox que no se expresan están agrupados en el núcleo en los llamados *Polycomb bodies*, e interaccionan entre ellos independientemente de la distancia, incluso de si son del complejo de Antennapedia o del Bithorax (Bantignies et al., 2011). Su activación no solo implica la pérdida de esta marca, sino también un aumento de H3K4me3, quedando separados los genes que están activos de los que se mantienen inactivos en diferentes estructuras topológicas (Lanzuolo et al., 2007).

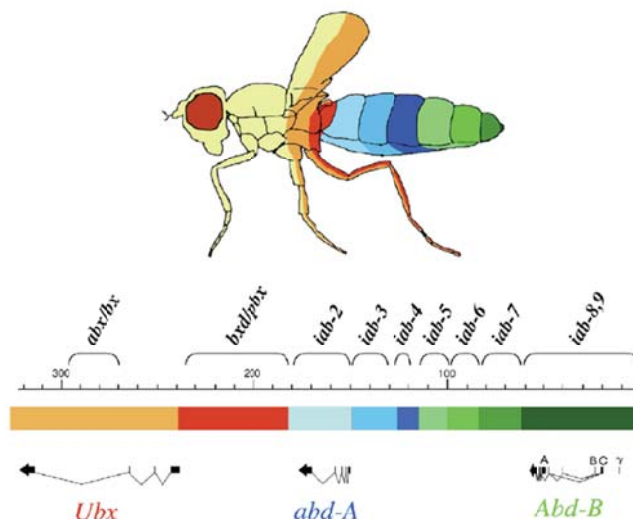
### **2.1.2. El complejo Bithorax**

El conjunto de genes homeóticos mejor estudiado es el complejo Bithorax, que se sitúa en el cromosoma 3R y tiene un tamaño de más de 300kb. Está formado por 3 genes: *Ultrabithorax (Ubx)*, *abdominal-A (abd-A)* y *Abdominal-B* (Figura 14), que son los encargados de controlar la identidad de cada uno de los segmentos que forman los dos tercios posteriores de la mosca. Se expresan de manera aditiva en el eje antero-posterior del embrión, lo que implica que en los segmentos posteriores se expresan los tres (Lewis, 1978, Celniker et al., 1989). *Ubx* se empieza a expresar en los segmentos posteriores del tórax, *abd-A* en los primeros segmentos del abdomen y *Abd-B* en la región trasera del abdomen. Están regulados por nueve dominios puestos en el mismo orden que los segmentos que controlan. Los dominios *abx/bx* y *bxl/pbx* son responsables de la expresión de *Ubx* en los parasegmentos 5 y 6 respectivamente, que corresponden a los segmentos T3 y A1 en el adulto. Los dominios de *iab-2* a *iab-4* regulan *abd-A* en los parasegmentos de 7 a 9, correspondiendo a los



## Introducción

segmentos A2, A3 y A4. Finalmente los dominios de *iab-5* a *iab-8,9* regulan la expresión de *Abd-B* en los parasegmentos 10-14, que corresponden a los segmentos abdominales A5-A9 (Boulet et al., 1991, Maeda and Karch, 2006).



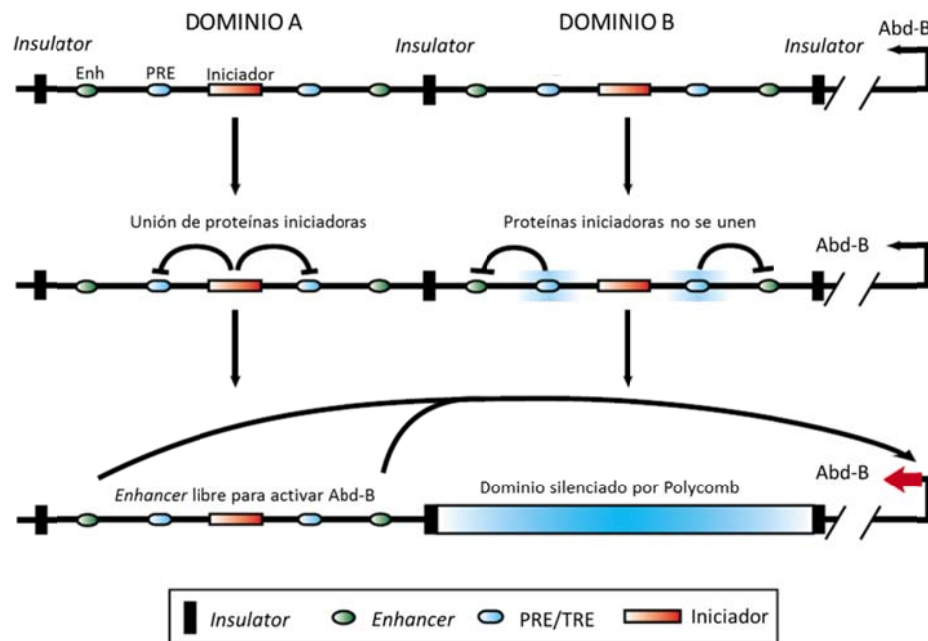
**Figura 14. El complejo Bithorax y los segmentos que controla.** Representación de los tres genes homeóticos de BX-C, *Ubx*, *abd-A* y *Abd-B* que determinan la parte posterior del organismo. A nivel genómico, las nueve regiones reguladoras (señaladas con arcos) que controlan su expresión están ordenadas colinealmente. En el dibujo, cada una de ellas se simboliza con el mismo color que el segmento que determina (Maeda and Karch, 2006).

### 2.1.2.1. Organización y funcionamiento de la región reguladora de *Abdominal-B*

La regulación de *Abd-B* está bajo el control de cuatro dominios, denominados infra-abdominales (*iab*), separados entre ellos por las regiones Front Abdominales (*Fab*), *Mcp*, *Fab6*, *Fab7* y *Fab8* (Karch et al., 1994, Cavalli and Paro, 1998, Barges et al., 2000, Perez-Lluch et al., 2008), para que el estado activo o silenciado de un dominio no se extienda a los adyacentes, actuando de barrera. En efecto, son *insulators*, y su correcto funcionamiento es crucial para que no haya errores de identidad entre segmentos. Los *Fabs* delimitan cada uno de los dominios, de manera que un *enhancer* no puede actuar en otro parasegmento.

El Modelo de los Dominios (Figura 15) (Mihaly et al., 2006) recoge todo lo que se conoce sobre el funcionamiento de la región reguladora de *Abd-B*. En éste se propone que cada dominio contiene todos los elementos necesarios para controlar la expresión, que comienza al inicio del desarrollo como una respuesta a las proteínas Gap y pair-rule, que se unen a elementos iniciadores en cada región del BX-C para determinar el patrón de expresión, y transmiten la señal de activación al resto de elementos del dominio (Shimell et al., 1994). Entre ellos, están los elementos de mantenimiento, que se encargan de mantener esta señal cuando los niveles de las proteínas que lo inician bajan. Éstos son PREs/TREs, que mantienen el estado reprimido o activo a lo largo de las divisiones, permitiendo que la identidad del segmento se mantenga inmutable durante el

desarrollo (Duncan, 1982, Dejardin et al., 2005). Los Fabs, actuando como *insulators*, evitan la extensión de la actividad o silenciamiento a dominios adyacentes.



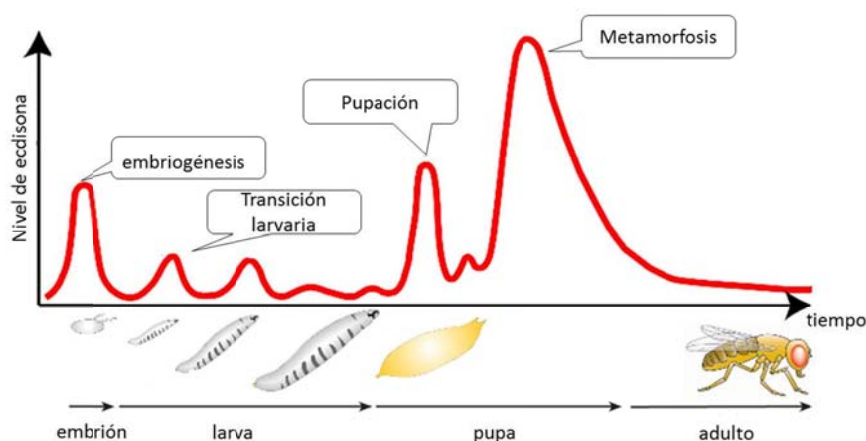
**Figura 15. Modelo de los Dominios.** Esquema representativo de cómo funciona cada dominio. En la parte superior se observan dos dominios, en el que en cada uno hay *enhancers*, PREs y un elemento iniciador. A la derecha está el gen diana *Abd-B*. La sombra azul simboliza el silenciamiento mediado por Polycomb. El ejemplo de la izquierda muestra la activación de un dominio, mientras que el de la derecha muestra uno no activado (Modificado de Mihaly et al., 2006).

Sin embargo este modelo supone que entre *enhancers* y promotor puede haber uno o varios Fabs ¿Cómo logran los *enhancers* hacer su función entonces? Uno de los modelos afirma que la manera de conseguirlo es gracias a las *promoter targeting sequence* (PTS), que se encuentran a un lado de los Fabs, y que hacen que estos *insulators* en lugar de tener un efecto a larga distancia, ejerzan un papel más local, delimitando y separando dominios contiguos (Zhou and Levine, 1999, Chen et al., 2005). Los *enhancers* serían capaces de contactar con el promotor sobrepasando el efecto bloqueante de los *insulators*. Los PTSs también contribuirían a la correcta expresión de los genes, para que cada *enhancer* de un segmento se ponga en contacto con el promotor independientemente de la actividad transcripcional de otros iab (Lin et al., 2010). Otro de los modelos propone, de acuerdo a la función antes descrita de los *insulators* como reguladores transcripcionales, que no solo aíslan cada iab del resto para que no se vean afectados por efectos activadores o silenciadores externos, sino que también, como se ha visto, organizan la estructura de la cromatina de toda la región reguladora permitiendo la expresión o silenciamiento (apartado 1.2.2.3.) (Cleard et al., 2006, Mihaly et al., 2006). Cualquiera de estos modelos propone los *loops* de cromatina como mecanismo de acción.

## 2.2. Regulación temporal de la esteroidogénesis

### 2.2.1. Transiciones entre estadíos larvarios y metamorfosis

Las transiciones morfológicas son otro de los procesos en los que la regulación de la expresión de genes cobra especial importancia. En insectos, estas transiciones incluyen los cambios entre estadíos larvarios y la metamorfosis, que tienen lugar en momentos muy concretos y dependen de los pulsos de las hormonas esteroideas, principalmente ecdisona (Yamanaka et al., 2013). Durante el desarrollo larvario sus niveles se incrementan de manera pulsátil controlando la transición de un estadío a otro. Al final del tercer estadío se produce un fuerte pulso que induce la pupación y comienzo de la metamorfosis. El proceso finaliza con otro pulso que marca la diferenciación de las estructuras del adulto (Figura 16). Este preciso control sugiere que la síntesis y/o degradación de los ecdisteroides deben ser finamente reguladas.



**Figura 16. Niveles de ecdisona durante el desarrollo.** Representación de los niveles de ecdisona que se producen a lo largo de las transiciones larvarias y metamorfosis, desde embrión hasta adulto (Modificado de [www.tsukuba.ac.jp](http://www.tsukuba.ac.jp)).

La ecdisona es producida durante las fases de larva y pupa en la glándula protorácica (PG), y es secretada a la hemolinfa donde se transforma en su forma activa, 20-hidroxiecdisona (20E), y se dirige a los tejidos donde va a desarrollar su acción. Aquí interacciona con su receptor (EcR) induciendo la activación de sus genes diana como *Broad (Br)*, *E74* o *E75* (Moeller et al., 2013).

El proceso es iniciado por el neuropéptido *protoracicotropic hormone* (PTTH) que es producido por dos pares de células neurosecretoras del cerebro de los insectos, y es liberado, actuando como señal para la PG (McBrayer et al., 2007). PTTH, a través de la activación de su receptor Torso y la cascada de *mitogen-activated protein kinase* (MAPK) (Rewitz et al., 2009), lleva a un aumento de la transcripción de las enzimas responsables de la síntesis de ecdisona, la familia de los genes *Halloween* (Gilbert, 2004). Éstos son los encargados de la producción de la hormona a partir de colesterol de los alimentos o fitosteroles, ya que los artrópodos no son capaces de sintetizar colesterol. En los últimos años ha ido habiendo más evidencias de que otros factores además de

PTTH, especialmente señales nutricionales, controlan la ecdisteroidogénesis actuando en la glándula protorácica (Layalle et al., 2008).

### **2.2.2. Genes *Halloween***

La mayoría de las enzimas que se engloban dentro de los genes *Halloween* pertenecen a la familia de las citocromo P450. Para muchas proteínas de esta familia no se ha conocido aún su función, pero varias se han visto implicadas en la biosíntesis de la ecdisona. Las que se conocen que participan en este proceso son *spookier* (*spok*) (Ono et al., 2006), *disembodied* (*dib*), *shadow* (*sad*) (Warren et al., 2002) y *phantom* (*phm*) (Warren et al., 2004). Además de las CYP450 hay otras enzimas que también contribuyen en la síntesis de la hormona tal como, *neverland* (*nvd*) (Yoshiyama et al., 2006), proteína *oxygenase-like*, *noppera-bo* (*nobo*) (Enya et al., 2014) que pertenece a la familia *glutathione S-transferase* y *shroud* (*sro*) (Niwa et al., 2010), que forma parte de las proteínas de *short chain dehydrogenase/Reductase*.

La regulación de la expresión de estos genes se conoce solo parcialmente. Se sabe que los ecdisteroides están implicados en *feedbacks* para regular su propia síntesis (Gilbert et al., 2002), pero además se ha visto que, en *Drosophila*, la afectación de algunos genes, por ejemplo *ecdysoneless* (*ecd*), *giant* (*gt*), *lethal(3)*dre4** (*dre4*) o *without children* (*woc*), llevan a bajos niveles de algunos ecdisteroides (Garen et al., 1977, Schwartz et al., 1984, Sliter and Gilbert, 1992, Wismar et al., 2000), sugiriendo que la regulación transcripcional de los genes implicados en la vía es clave en el proceso.

Diversos estudios han confirmado que los reguladores transcripcionales tienen un efecto en la producción de ecdisona. El primer ensayo fue realizado en *Drosophila* (Parvy et al., 2005), y se investigó la contribución de *Dmhfz-f1* (proteína relacionada con la metamorfosis y que es ortóloga de la proteína de vertebrados *Steroidogenic Factor 1* (SF1) que favorece la transcripción de genes relacionados directa o indirectamente con la síntesis esteroidea (Val et al., 2003)). Demostraron que la falta de la proteína producía una disminución de la transcripción de los genes *dib* y *phm* en los momentos en que se debieran expresar. Esto confirmaba que algunos factores transcripcionales, y no solo los *feedbacks* producidos por las propias hormonas, estaban involucrados en esta regulación. Sin embargo, la disminución de la cantidad de hormona sintetizada era mucho mayor que la esperada en función de la bajada de los niveles de expresión de estos genes, lo que sugería que el efecto producido por la disminución de la expresión génica era solo una parte del circuito regulatorio que controlaba los niveles hormonales, y que otros aspectos debían influir en el proceso.

Este estudio pionero no explicaba el mecanismo por el que los factores transcripcionales regulan la expresión de los genes *Halloween*. Años más tarde se caracterizaron los elementos reguladores necesarios para la expresión de *phm* y *dib* en la PG de *Drosophila*, incluyendo el promotor de *phm* así como el *enhancer* de *dib*. Se demostró que el factor transcripcional Broad (Br-Z4) aumentaba la producción ecdisteroidogénica a través de la interacción con estos elementos (Moeller et al., 2013).

## **Introducción**

Se ha continuado investigando la participación de factores de transcripción en la biosíntesis de ecdisona, y un análisis reciente ha demostrado que la proteína *Ventral veins lacking* (Vvl) y el receptor nuclear *Knirps* (Kni) desempeñan funciones esenciales en la PG durante el desarrollo larvario (Danielsen et al., 2014). Se ha visto que, en la PG, ambas son capaces de unirse a elementos que favorecen la expresión de genes implicados en la síntesis de ecdisona como *phm*, *dib* y *spok* para controlar de esta manera su expresión. Sin embargo, la falta de estas proteínas no implica la disminución de la expresión de otros genes de la vía como *nvd*. También han mostrado que el regulador transcripcional *Molting defective* (Mld) controla los primeros pasos de la vía de síntesis, aunque no han encontrado secuencias para su unión en la región reguladora de *spok*, por lo que no detallan el mecanismo de acción. La ausencia de las proteínas en estudio específicamente en la PG produce un retraso en el desarrollo, consistente con la afectación en la producción de la hormona, como comprueban al suplementar con la forma activa de la ecdisona. Además, han demostrado que son necesarias para mantener las señales transmitidas a la PG por las vías de insulina/TOR o PTTH, concluyendo que tanto Vvl como Kni pueden regular la vía de síntesis de la ecdisona de manera directa (a través de los genes *Halloween*) e indirecta (por estas dos vías que acabamos de mencionar).

En los últimos años se ha avanzado mucho en este campo, pero son necesarios más estudios sobre la regulación de la vía de biosíntesis de la ecdisona y de los factores implicados en ello.

## **3. HASPIN**

Haspin, cuyo nombre proviene de *haploid germ cell-specific nuclear protein kinase*, fue descrita por primera vez por Nishimune y sus colaboradores (Tanaka et al., 1999) como una proteína con capacidad de unión al DNA, actividad kinasa sobre serinas y treoninas e implicada en la diferenciación celular.

Su secuencia de aminoácidos muestra varios motivos conocidos (Tanaka, Yoshimura et al. 1999) como el de las proteínas con actividad kinasa (aunque solo tiene 3 de los 12 subdominios kinasa conservados) y contiene muchos sitios susceptibles de ser dianas de proteínas kinasas.

Según la base de datos OrthoDB, se trata de una proteína conservada en diferentes especies. Su dominio C-terminal, el que posee la actividad kinasa, tiene gran homología entre *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae* o mamíferos, como ratón o humanos (Higgins, 2001). Sin embargo, el dominio N-terminal, para el reconocimiento de los sustratos diana, presenta gran variabilidad.

La descripción de la localización y la función que desempeña Haspin en células de mamíferos fue descrita por Dai en 2005. Se observó que no solo estaba presente en testis como se había sugerido inicialmente (Tanaka et al., 1999), sino que se encontraba en diversos tipos celulares (Higgins, 2010). Mediante fraccionamiento se detectó su presencia nuclear en interfase así como su asociación con los cromosomas durante mitosis. A través de la proteína tagueada se observó que estaba presente en los cromosomas mitóticos desde profase hasta anafase. En metafase se encontraba unida a los brazos de los cromosomas, pero especialmente a los centrómeros. También se vio que localizaba en centrosomas desde prometafase hasta telofase (Dai et al., 2005).

Además de fosforilarse a sí misma, hasta ahora solo se conoce un sustrato de esta kinasa, la treonina en posición 3 de la histona H3, H3T3ph, aunque recientemente se han propuesto otros posibles sustratos de la proteína (Maiolica et al., 2014). La fosforilación de la propia proteína solo se produce durante la mitosis. A pesar de la poca detección de H3T3 durante interfase, se observa en células que están en la última fase de G2 con un patrón puntuado, uniéndose a los brazos de los cromosomas. En mitosis se concentra también aquí, en centrosomas y en el *spindle*, pero sobre todo en regiones centroméricas (Dai et al., 2005).

La actividad de Haspin está mucho más incrementada durante mitosis, hecho que se explica porque la proteína presenta un dominio básico inhibitorio (HBIS) que es capaz de autoinhibir el dominio kinasa en *Xenopus* y células humanas. Su función se ve contrarrestada por la fosforilación por parte de Polo del dominio N-terminal justo antes de entrar en mitosis (Ghenoiu et al., 2013). Sin embargo, se ha visto que esta autoinhibición no es constante, sino que sigue un mecanismo ON-OFF, que permite la fosforilación de H3T3 en interfase, aunque de manera más lenta a como lo hace en mitosis. De hecho, desde el comienzo de su estudio se propuso que su dominio kinasa estaba siempre activo ya que su sitio de activación presenta una composición característica que provoca que sea constitutivamente activo (Higgins, 2001, Villa et al., 2009, Higgins, 2010). Además, esto fue demostrado porque la sobreexpresión de la proteína generaba la fosforilación de H3T3 durante todo el ciclo celular (Dai et al., 2005).

En un estudio con células humanas se observó que la carencia de Haspin impedía la correcta alineación de los cromosomas en mitosis, produciéndose una acumulación de células en prometafase y una disminución en anafase y telofase. La sobreexpresión de la proteína no afectaba a la progresión del ciclo en interfase ni a la entrada en mitosis, pero sí a su progresión, como se detectó por la acumulación de células en profase/prometafase y la disminución en anafase/telofase, con el consiguiente retraso para entrar de nuevo en G1. Estos resultados señalaban la importancia de mantener los niveles de Haspin dentro de unos límites para la correcta progresión del ciclo celular (Dai et al., 2005). Posteriormente se afirmó que Haspin era necesaria para la localización de Aurora B en centrómeros (Yamagishi et al., 2010). Esta proteína controla la correcta alineación de los cromosomas metafásicos a través de la unión kinetocoro-microtúbulos, y la progresión hacia anafase a través del *Spindle Assembly Checkpoint* (SAC). El efecto que produce la falta de Haspin es como el que causa la carencia de Aurora B. Ambas proteínas no solo son necesarias para la correcta posición de los cromosomas metafásicos sino también para el correcto funcionamiento del SAC. Una vez

## **Introducción**

alcanzada la metafase, en presencia de inhibidores de Haspin o Aurora B se producía una salida temprana de mitosis incluso cuando se provocaba químicamente que los microtúbulos estuvieran fuertemente alterados. Todos estos resultados sugerían que Haspin regulaba la progresión del ciclo a través de su acción sobre Aurora B (Wang et al., 2012).

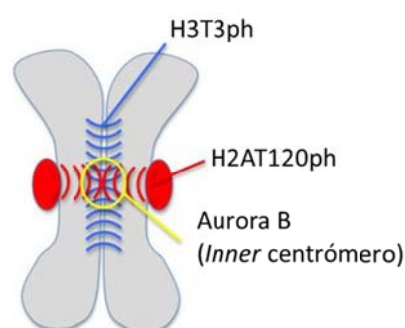
El efecto de Haspin sobre la alineación de los cromosomas podría ser debido, al menos en parte, a una pérdida prematura de la cohesión entre las cromátidas hermanas. Esta observación hizo que se propusiera una relación de Haspin con el complejo de las cohesinas, que son las responsable de ello. Se detectó que la falta de Haspin provocaba una disminución de cohesinas en los cromosomas, lo que desembocaba en la separación de las cromátidas (Dai et al., 2006). De hecho, el defecto en la alineación de los cromosomas que produce la falta de Haspin es similar al que conlleva la depleción de la cohesina Scc1 (ortólogo de Rad21 de *Drosophila*) (Dai et al., 2006). Además, la sobreexpresión de Haspin impide la disociación de las cohesinas de los brazos de los cromosomas, por lo que parece evidente la existencia de una relación entre Haspin y el complejo. A pesar de que no hay evidencias directas, se ha sugerido que Haspin no afecta al posicionamiento de las cohesinas, sino más bien a su mantenimiento una vez unidas al DNA, probablemente a través de la interacción con pds5. Yamagishi describió en 2010 que pds5 interactúa con la región N-terminal de Hrk1 (ortólogo de Haspin en *S. pombe*), y que esta proteína es necesaria para la localización de Hrk1 y de la H3T3ph por todo el cromosoma, pero que su propia localización no depende de Hrk1. Además, vio que la localización de pds5-Hrk1-cohesinas coincidía en los sitios del cromosoma analizados (Yamagishi et al., 2010). La necesidad de PDS5 para la localización de H3T3ph fue también confirmada para mamíferos (Carretero et al., 2013).

### **3.1. Haspin y la segregación del material genético**

#### **3.1.1. Mitosis**

El principal objetivo de la división celular es la transmisión equitativa del material genético duplicado a las dos nuevas células hijas. El *Cromosomal Passenger Complex* (CPC), formado por Borealin, Survivin, *inner centromere protein* (INCENP) y Aurora B kinasa tiene un papel central en la correcta segregación de los cromosomas. De hecho, la alteración de su función produce defectos en este proceso por desestabilización de la unión kinetocoro-microtúbulos, fallo en la función del SAC y en la formación del *spindle*. La regulación de la localización de CPC es clave para que el proceso se desarrolle de manera adecuada. En todos los organismos modelo estudiados, Aurora B se localiza en los *inner centrómeros* durante prometafase y metafase. Además, el hecho de que el modo de regulación del reclutamiento de CPC a esta localización esté conservado evolutivamente implica que la acumulación centromérica de CPC es crucial para la función de Aurora B (van der Horst and Lens, 2014).

Se ha observado que Haspin juega un papel fundamental en la localización de CPC en los centrómeros. Este posicionamiento de CPC depende de Survivin, que interacciona con H3T3ph, marca realizada por Haspin, y de Borealin, que une H2AT120ph, modificación mediada por Bub1, a través de proteínas Shugoshin (Sgo1). El complejo se localiza entre estas dos fosforilaciones, lugar denominado *inner centrómeros*, y da una explicación de porqué CPC se encuentra ahí (Figura 17). Se ha confirmado que ambas modificaciones de histonas son determinantes para ello, ya que se ha observado que la falta de Bub1 conlleva una distribución de CPC por la cromatina y que la de Haspin provoca que Aurora B siguiera unida a los brazos de los cromosomas (Boyarchuk et al., 2007, Yamagishi et al., 2010). De aquí se deduce que reguladores de la localización y de la actividad de Haspin y Bub1 también repercuten en CPC. Las señales que determinan el posicionamiento de CPC son a su vez controladas por Aurora B, sugiriendo la existencia de diversos mecanismos de *feedback*.



**Figura 17. Localización de Aurora B en *inner centrómeros* depende de Haspin y Bub1.** Esquema representativo de la localización de H3T3ph realizada por Haspin y H2AT120ph por Bub1, cuyo cruce define el inner centrómero, donde se localiza Aurora B.

En lo que respecta a Haspin, se ha visto que esta proteína es fosforilada en múltiples sitios por Aurora B para controlar su actividad kinasa sobre H3T3 (Wang et al., 2011), por lo que Aurora B refuerza su propia localización a través de la regulación de Haspin. Además, se ha descrito que Aurora B es capaz de activar Plk1, kinasa necesaria para fosforilar el dominio autoinhibitorio de Haspin desencadenando la reactivación del dominio kinasa (Carmena et al., 2012). Por otro lado, H3T3 es desfosforilada por la fosfatasa PP1 $\gamma$ . Repo-Man la recluta, eliminando H3T3ph de los brazos de los cromosomas, y permitiendo la liberación de CPC de los mismos. Aurora B fosforila esta enzima en prometáfase y metafase en centrómeros impidiendo su acción y, por tanto, permitiendo la presencia de H3T3ph y, por consiguiente la acumulación de CPC. Repo-Man también interacciona con la fosfatasa PP2A que elimina la fosforilación causada por Aurora B sobre Repo-Man, permitiendo su posterior eliminación (Qian et al., 2011, Qian et al., 2013). Estos procesos implican un *feedback* entre Haspin y Aurora B y entre Aurora B y las fosfatasas que determinan los niveles de H3T3ph en centrómeros y, por tanto, la acumulación centromérica de CPC. Por otro lado, la localización de Haspin también afectará a los niveles de fosforilación de H3T3 en los centrómeros.

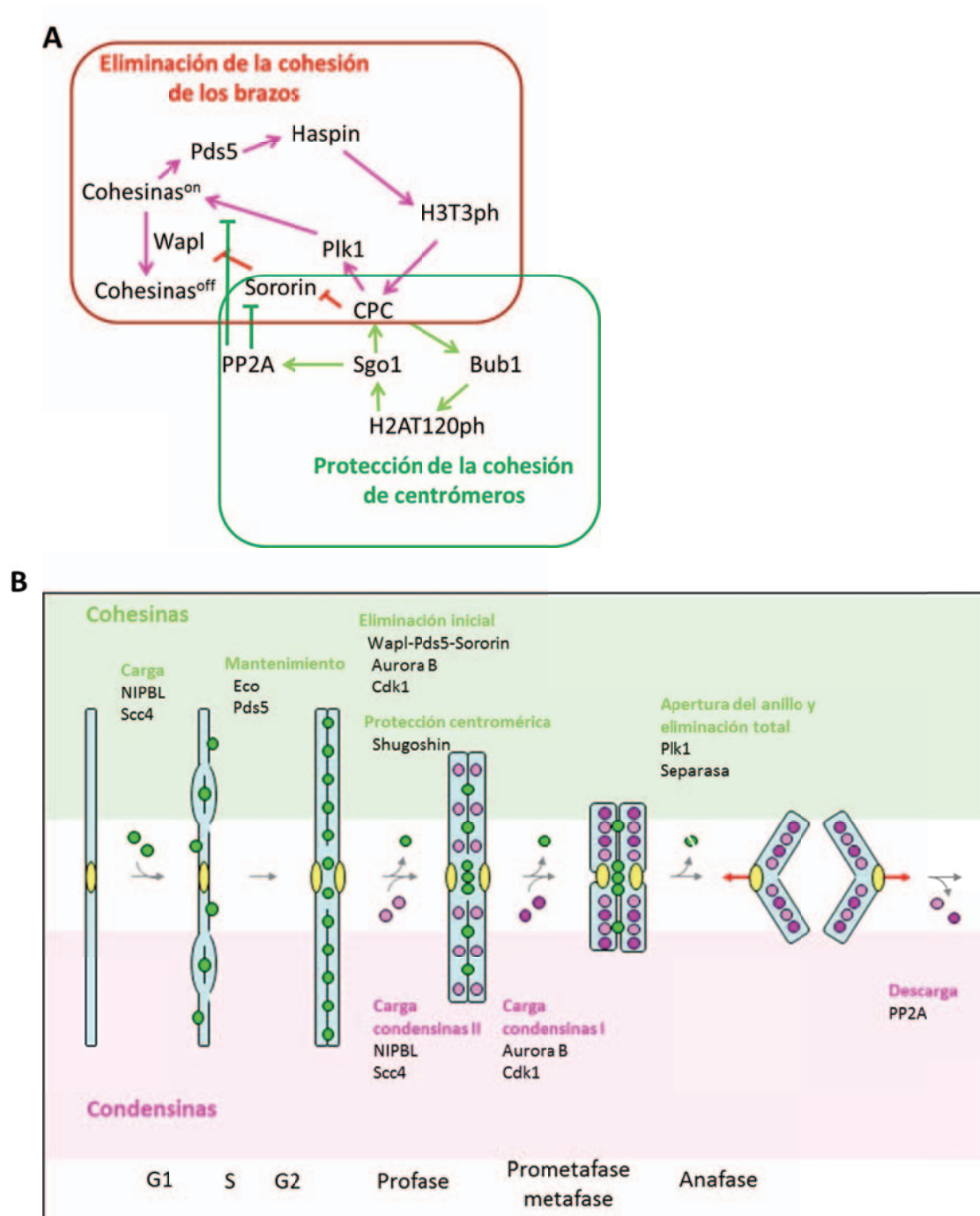


## **Introducción**

### **3.1.1.1. Dinámica de unión de cohesinas y condensinas al DNA**

Las cohesinas son cargadas a los brazos de los cromosomas por NIPBL y Scc4 en G1, existiendo un equilibrio entre su mantenimiento (mediado por Pds5 y Eco) y su separación (controlada por Pds5 y Wapl). Durante las fases S y G2 la unión a los brazos de los cromosomas se estabiliza gracias a Eco y Pds5 para conseguir la cohesión de las cromátidas hermanas. La liberación de las cohesinas de los brazos en profase está controlada por tres factores: Pds5 (Pds5 A/B en mamíferos), que une directamente las cohesinas; Wapl que puede eliminar las cohesinas unidas a la cromatina abriendo el anillo e interactuando con Pds5; y Sororin, que compite con Wapl por los sitios de unión en Pds5, protegiendo de esta manera las cohesinas unidas. Aurora B y Cdk1 fosforilan Sororin en múltiples sitios, disminuyendo su interacción con Pds5, permitiendo así a Wapl unirse a las cohesinas de los brazos de los cromosomas y eliminarlas. Además, Plk1 puede fosforilar directamente el complejo y promover la eliminación de las cohesinas. Mientras que Wapl es reclutada a los brazos de los cromosomas, las cohesinas centroméricas están protegidas por Shugosin. El hecho de que Pds5 reclute Haspin sugiere que el complejo de las cohesinas recluta indirectamente CPC, que mediante la fosforilación de Sororin favorece la eliminación de las cohesinas de los brazos. Sin embargo, CPC recluta a su vez Bub1 que fosforila H2AT120, marca reconocida por Sgo1, el cual protege la cohesión centromérica a través del reclutamiento de la fosfatasa PP2A, que elimina las fosforilaciones mediadas por Aurora B y Plk1 para mantener la cohesión. La total liberación de las cohesinas en la anafase se produce mediante la fosforilación de SA a través de Plk1 y por la proteasa separasa, que en anafase es capaz de interactuar con Rad21 para alterar el anillo y permitir la liberación de las cromátidas hermanas (Figura 18) (Higgins, 2010, Trivedi and Stukenberg, 2015).

En los últimos años, un gran cuerpo de evidencias constata la acción no solo de cohesinas sino también del complejo de condensinas en la regulación de la correcta segregación de los cromosomas. En el momento en que empieza la liberación de las cohesinas de los brazos de los cromosomas, comienza la carga progresiva de condensinas, uniéndose primero condensinas II por NIPBL y Scc4. Mientras que las condensinas II tienen una localización nuclear a lo largo de todo el ciclo celular, las condensinas I son citoplasmáticas en interfase, y entran en el núcleo tras la ruptura de la envoltura. CPC puede promover la compactación de los cromosomas cargando condensinas I a los brazos. Aurora B y Cdk1 las unen al DNA, lo que culmina con la formación de los cromosomas metafásicos. Este proceso, la condensación de los cromosomas, es un prerrequisito para la separación final de las cromátidas, que se consigue mediante la total liberación de cohesinas. Al final de la anafase la inhibición de Cdk1 impide la asociación de condensinas con la cromatina, lo que promueve la descondensación para permitir a las células volver a interfase. La fosfatasa PP2A también se ha visto implicada en la defosforilación de condensinas (Figura 18B) (Higgins, 2010).



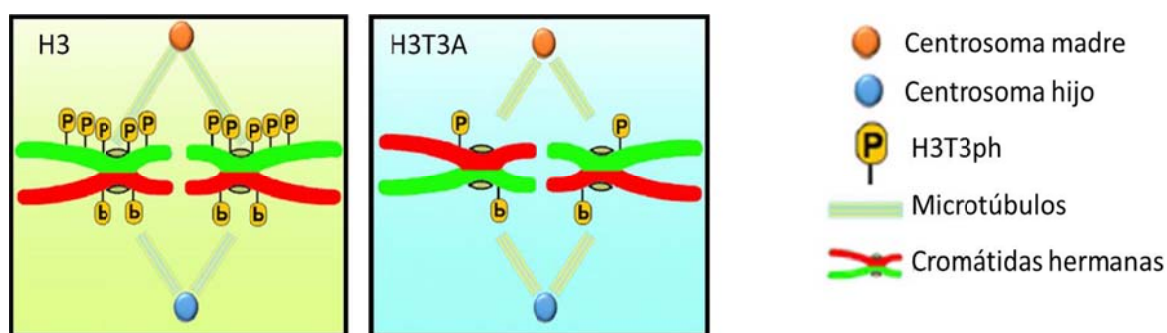
**Figura 18. Regulación de la segregación mitótica.** A/ Esquema representativo de la red de señalización que regula la cohesión de las cromátidas hermanas durante mitosis. La parte responsable de la eliminación de cohesinas en brazos está representada en rojo, y la de protección centromérica en verde (Modificado de Trivedi and Stukenberg, 2015). B/ Representación de los pasos que siguen cohesinas y condensinas en su carga y posterior separación del DNA. Se indica el proceso que tiene lugar en cada fase del ciclo celular, así como las proteínas responsables de desempeñar estas funciones (Modificado de <http://www.riken.jp/chromdyna/en/research/project02.html>).

## Introducción

### 3.1.1.2. Segregación cromosómica en células pluripotenciales

La segregación del material genético se produce de una manera especial en la división celular de muchos tipos de células pluripotenciales de adulto, como es el caso de células germinales (GSC, *germline stem cell*) de machos de *Drosophila*. En estas células la segregación se hace de forma asimétrica. Antes de la mitosis, una vez que el material genético ha sido replicado, la histona 3 preexistente y la recién sintetizada se distribuyen diferencialmente en las dos cromátidas hermanas, y durante la mitosis, la célula hija que será una nueva célula pluripotencial recibirá la histona 3 que previamente ya existía mientras que la que dará lugar a la célula diferenciada o gonioblasto obtendrá la recién sintetizada (Tran et al., 2012).

Haspin juega un papel fundamental en esta segregación asimétrica. De hecho, sería la fosforilación en la treonina 3 de la histona 3 lo que distinguiría la preexistente de la sintetizada (Figura 19). Tanto el defecto (estudiado gracias a un transgén en el que se ha convertido la treonina en alanina, que no se puede fosforilar) como la presencia constante (analizada con un transgén fosfomimético en que la treonina es sustituida por aspartato) de esta PTM provocan alteraciones en este proceso; las células hijas presentan un patrón de herencia bastante simétrica e incluso un patrón de asimetría invertida. Además, las células hijas sufren una pérdida de capacidad pluripotencial y presentan fenotipos tumorales. El hecho de observar una distribución y fenotipos similares en ambos casos sugiere que el momento en que se produce la fosforilación de H3T3 es importante para la correcta actividad de GSC. Por otro lado, el fenotipo que causa la ausencia de Haspin refuerza el que produce el transgén no fosforilable, lo que sugiere que Haspin puede tener otras dianas aún desconocidas además de H3T3ph en GSC en *Drosophila* (Xie et al., 2015).



**Figura 19. H3T3ph es necesaria para la segregación asimétrica del material genético.** Esquema representativo de la división asimétrica en células pluripotenciales. En la parte de la izquierda se observa cómo la histona 3 preexistente y la recién sintetizada se distribuyen diferencialmente en las dos cromátidas hermanas por la presencia de H3T3ph. Si la treonina en posición 3 no puede ser fosforilada (derecha), la segregación no se produce de manera asimétrica. H3T3A: residuo no fosforilable en el que se ha cambiado treonina por alanina (Modificado de Xie et al., 2015).

### **3.1.2. Meiosis**

Haspin, tal como hemos descrito para mitosis, también actúa como regulador del CPC y de la segregación cromosómica durante meiosis. Sin embargo, cabe señalar diferencias importantes en su participación en el proceso (Nguyen et al., 2014). Ya que meiosis I implica mantener las cromátidas hermanas juntas mientras que se separan los cromosomas homólogos, es posible que Haspin juegue papeles adicionales o diferentes en la regulación de la meiosis I de lo que hace regulando la segregación de cromosomas en mitosis.

Las células germinales contienen un homólogo de Aurora B llamado Aurora C, que también forma parte del CPC (Li et al., 2004, Sasai et al., 2004). En meiosis, Haspin también recluta al CPC para ejercer su función para regular la alineación de los cromosomas y la unión kinetocoro-microtúbulos, así como la correcta actividad del SAC. No obstante, mientras que en mitosis la fosforilación de H3T3 localiza en centrómeros, durante metafase I la histona modificada está presente no solo aquí sino también por todo el eje entre las cromátidas (ICA, *interchromatid axis*). Esta posición, específica de meiosis, es un requisito indispensable para que CPC se sitúe en él, y pueda controlar la correcta unión kinetocoro-microtúbulos. De hecho, la ausencia de Haspin provoca la falta de CPC en este lugar. Esta deslocalización altera la capacidad de los oocitos para corregir la unión incorrecta de kinetocoro-microtúbulo, dando desalineación de los cromosomas en metafase I y aneuploidía en metafase II (Nguyen et al., 2014).

Haspin también es necesaria para la correcta progresión del ciclo celular en meiosis, de manera similar a como hemos señalado en células somáticas. La falta de Haspin provoca un retraso en el comienzo de la meiosis. Aunque finalmente se produce, las células son arrestadas en metafase I, ya que también la progresión hacia metafase II se ve retrasada. Por otro lado, mediante la eliminación de Haspin una vez alcanzada la metafase, se ha observado que también es necesaria para el correcto funcionamiento del SAC en células germinales, ya que las células desprovistas de Haspin entran en anafase I de manera prematura a pesar de que los cromosomas no estén correctamente posicionados o la unión kinetocoro-microtúbulo no sea la correcta. De hecho, en ausencia de Haspin, MAD1, componente principal de SAC, no es capaz de unirse a los kinetocoros para bloquear la evolución del ciclo cuando la unión kinetocoro-microtúbulo no es la adecuada (Wang et al., 2015).

#### **3.1.2.1. Dinámica de unión de cohesinas y condensinas al DNA**

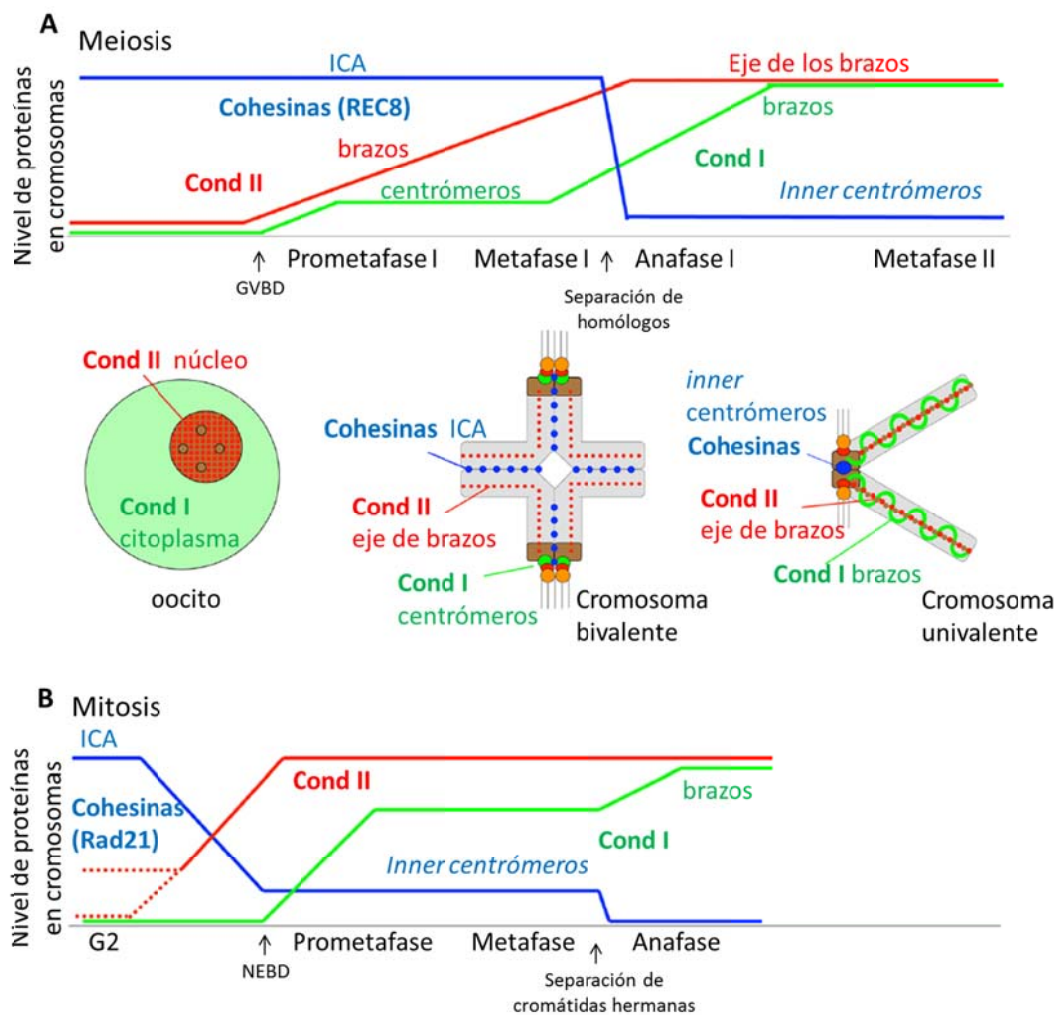
Como en mitosis, en meiosis también existe la relación cohesinas-condensinas. En este caso, es más evidente la relación de Haspin con condensinas, y no tan clara con cohesinas. La perturbación de la actividad kinasa de Haspin genera anomalías en la estructura de los cromosomas (Nguyen et al., 2014); su ausencia provoca fenotipos de cromosomas redondeados y compactos, mientras que su sobreexpresión conlleva el efecto contrario, que los cromosomas aparezcan descompactados y extendidos. Por tanto, la condensación de los cromosomas se ve afectada por Haspin (Wang et al., 2015). De hecho, Haspin regula la carga de condensinas en cromosomas meióticos. Se ha visto que

## **Introducción**

SMC2 (componente común de los dos complejos de condensinas) comparte la localización con Haspin en centrómeros y en los ejes de las cromátidas hermanas en metafase I (Lee et al., 2011), y la alteración de Haspin afecta a la presencia de SMC2 en los cromosomas de metafase I (Nguyen et al., 2014). Sin embargo, condensinas I y condensinas II siguen diferente distribución durante meiosis (Figura 20A). El primer grupo solo se aprecia en el núcleo a partir de metafase I en regiones próximas a los centrómeros, y solo comienza a estar presente en los brazos desde anafase I, donde perdura hasta metafase II. En cambio, las condensinas II están presentes en el núcleo durante todo el ciclo y ya se detectan en los brazos desde metafase I, donde permanecen en el eje central de las cromátidas hasta metafase II. Si comparamos su localización con cohesinas, REC8 (cohesina específica de meiosis) desde profase I se encuentra en ICA, por lo que en metafase I no comparte localización con condensinas, y se mantiene ahí hasta anafase I. A diferencia de lo que ocurre en mitosis, la cohesión entre cromátidas no se ve alterada por Haspin durante meiosis I como se ha visto por la presencia de REC8 en los brazos de los cromosomas durante metafase I en ausencia de Haspin, indicando que la cohesión entre cromátidas se mantiene intacta y protegida de una proteólisis prematura. Sin embargo, en metafase II se aprecian cromátidas hermanas separadas, sugiriendo que la protección del complejo tras metafase I se ve alterada por la falta de Haspin (Lee et al., 2011, Nguyen et al., 2014).

Las diferencias espacio-temporales entre ambos tipos de condensinas son acordes con las diferentes funciones que desempeñan; mientras que a las primeras se les ha asignado un papel en la correcta orientación de los cromosomas bivalentes, las condensinas II influyen en la condensación, y su falta provoca pérdida de compactación y rigidez, como se aprecia por su aspecto más extendido. Así, la morfología de los cromosomas que produce la alteración de Haspin es opuesta a la que produce la pérdida y ganancia de condensinas II (Lee et al., 2011).

Se ha visto un retraso en la carga de condensinas I y II a los cromosomas en meiosis respecto mitosis (Figuras 20A y 20B). ¿Qué es lo que lo provoca? La presencia simultánea en los cromosomas de cohesinas y condensinas es generalmente incompatible en ambos procesos, y una diferencia fundamental es la gran cantidad de cohesinas que permanecen entre las cromátidas hermanas desde prometafase I hasta metafase I en meiosis. Puede que su presencia se contraponga a la de las condensinas II, enlenteciendo su carga y acción. Respecto a condensinas I, el retraso es aún mayor, y solo pueden unirse a los brazos a partir de metafase I, una vez que las cohesinas ya no se encuentran ahí. A pesar de las diferencias en la dinámica de cohesinas y condensinas en mitosis y meiosis, se aprecian ciertas similitudes, como el orden de unión de los complejos de condensinas al DNA (primero las condensinas II y posteriormente las condensinas I) y que en ambos procesos cohesinas y condensinas II son parcialmente compatibles, mientras que cohesinas nunca coexisten con condensinas I, ni en mitosis ni meiosis.



**Figura 20. Dinámica de unión de cohesinas y condensinas.** A/ Esquema representativo de la dinámica de distribución de cohesinas (azul), condensinas I (verde) y condensinas II (rojo) en los cromosomas meióticos en función del nivel de proteína detectada, en la localización que se indica en cada fase del ciclo señalada. ICA: *interchromatid axis* GVBD: *germinal vesicle breakdown*. B/ Esquema representativo de la dinámica de distribución de cohesinas (azul), condensinas I (verde) y condensinas II (rojo) en los cromosomas mitóticos en función del nivel de proteína detectada, en la localización que se indica en cada fase del ciclo señalada. NEBD: *nuclear envelope breakdown* (Modificado de Lee et al., 2011).

Por tanto, la función de Haspin, cohesinas y condensinas en mitosis y meiosis se encuentra relacionada, sugiriendo que pueden también estarlo en interfase en la organización del genoma y la regulación de la expresión génica.



# Objetivos





Estamos interesados en los factores y mecanismos involucrados en la organización de la cromatina en diferentes dominios y su participación en la regulación de la expresión génica. En relación a esto, los objetivos de esta tesis son:

-Estudio de nuevas funciones de proteínas con actividad *insulator*: conocer el papel de CTCF en la esteroidogénesis en *Drosophila melanogaster*.

-Identificación de proteínas involucradas en la actividad *insulator* de CP190: analizar el complejo proteico de CP190, caracterizar sus componentes y estudiar la función de los mismos.

-Búsqueda de nuevas proteínas con actividad *insulator*: descubrir proteínas cuya actividad *insulator* no se conoce y profundizar en el modo en que actúan.



# Materiales y métodos



## 1. MATERIALES

### 1.1. Oligonucleótidos

Nombre	Secuencia 5' -> 3'	Experimento
RpL23 F	GACAACACCGGAGCCAAGAACC	PCR cuantitativa
RpL23 R	GTTTGCCTGCTGCCGAATAACCAC	PCR cuantitativa
EcR F	TGCGAAGAAGAGCAAGAAGG	PCR cuantitativa
EcR R	CAGGTGAGGGCGTTGTAGTG	PCR cuantitativa
E75B F	GTGGTGGGATGAAGCACGTA	PCR cuantitativa
E75B R	GAAGTGGTGGTCAGTAGGGC	PCR cuantitativa
spok F	TATCTCTTGGGCACACTCGCTG	PCR cuantitativa
spok R	GCCGAGCTAAATTTCTCCGCTT	PCR cuantitativa
dib F	TGCCCTCAATCCCTATCTGGTC	PCR cuantitativa
dib R	ACAGGGTCTTCACACCCATCTC	PCR cuantitativa
sad F	CCGCATTGAGCAGTCAGTGG	PCR cuantitativa
sad R	ACCTGCCGTGTACAAGGAGAG	PCR cuantitativa
nvd F	GGAAGCGTTGCTGACGACTGTG	PCR cuantitativa
nvd R	TAAAGCCGTCCACTTCCTGCGA	PCR cuantitativa
phm F	GGATTTCTTTGCGCGGATGTG	PCR cuantitativa
phm R	TGCCTCAGTATCGAAAAGCCGT	PCR cuantitativa
nobo F	TCAAGGGCGAGCAATTCCAA	PCR cuantitativa
nobo R	ACAACTACCGCCCTCATCG	PCR cuantitativa
sro F	AGCAGCTGAAGGTCGATAGC	PCR cuantitativa
sro R	TCCGACTTGATGTTGTGGCA	PCR cuantitativa
E74B F	CGCGAGTTCAAAGTGCTCTA	PCR cuantitativa
E74B R	GGAGGGAGAGTGGTGGTGT	PCR cuantitativa
haspin U	CTGGTCATCGTTTTACCCA	PCR cuantitativa
haspin L	CCTGTGAACCTTCGTATTGATGC	PCR cuantitativa
actina F	GCGCGGTTACTCTTTCACCA	PCR cuantitativa
actina R	ATGTCACGGACGATTCACG	PCR cuantitativa
tubulina F	ACCTGAACCGTCTGATTGGC	PCR cuantitativa
tubulina R	GCAGAGAGGCGGTAATCGAG	PCR cuantitativa
cDNA Haspin U	AAACTCGAGATGCTCAGCATTAGC	Clonaje cDNA de Haspin
cDNA Haspin L	AAACTCGAGTTAATTTAAATTAAACAG	Clonaje cDNA de Haspin
promHaspin_FORWARD	GGTACCAATGACAGCGGCAAATTCTC	Clonaje promotor endógeno de Haspin
promHaspin_REVERSE	CCTGTTGTGCCTCCYAATCG	Clonaje promotor endógeno de Haspin
seqTopo Haspin	CCGGGAAAATGGAGAAAATC	Comprobar cDNA de Haspin
pUASTattB-PRcDNA Rev	GCAGCTTTTTCCTTTGTGGT	Comprobar cDNA de Haspin

## **Materiales y métodos**

pAtt Upper	GGATTTCACTGGAAGTAGGC	Comprobar promotor endógeno de Haspin
Seq PROM-cDNA	GCAACAGTTGTGCAATTAACCT	Comprobar promotor endógeno de Haspin
Seq PROM-cDNA 2	ACCACTTCCGTTTAATACTGCA	Comprobar promotor endógeno de Haspin
MutKinH420A-F	ACCTTGCTTTGGGAAACATATTA	Clonaje mutación dominio kinasa
MutKinH420A-R	CCCTGTGCTCAAATTGATACTC	Clonaje mutación dominio kinasa
MutKinK282M-F	TTAATGATTATACCTTTGGAGGGTTC	Clonaje mutación dominio kinasa
seqMUTK282M	CGGGAAAATGGAGAAAATCGC	Comprobar mutación dominio kinasa
5Pout-2	GGCTATCGACGGGACCACC	Comprobar mutante nulo
3Pout	AGTGGATGTCTCTTGCCGAC	Comprobar mutante nulo
Screen 300-2000F	AGCTCAGGATCGAAGACAGT	Comprobar mutante nulo
Screen 300-2000R	ACCCCCACTCTAAAGGCCCTA	Comprobar mutante nulo
Screen 2000-300F	TCGGATCTACACTTGGCCTTC	Comprobar mutante nulo
Screen 2000-300R	CTGTTGTGCCTCCTAATCGTT	Comprobar mutante nulo
Screen1500F	TGAGGCGTGAAAAGTGGAGA	Comprobar mutante nulo
Screen1500R	CGGGTATCTGATTGTGCGGA	Comprobar mutante nulo
Screen 4000F	TGGAGACCAAGCCATGATCA	Comprobar mutante nulo
Screen 4000R	TCTCTCTCTTCGAGCACT	Comprobar mutante nulo
Screen 6000F	TGTAGGATGAGGGTTGCGTT	Comprobar mutante nulo
Screen 6000R	ACCAACGCCAGCATCTATA	Comprobar mutante nulo
SeqATGHaspin F2	ATTTAATTTAATGCTCAGC	Comprobar mutante nulo
MCP T	AGGCAGCCATCAATGTTGC	ChIP
MCP C	TACACATAATACATGGCCGCG	ChIP
83500 U Fab7	GCAGCCATCATGGATGTGAA	ChIP
83500 L Fab7	GGAATACCGCACTGTCGTAGG	ChIP
Fab8 C2	GCTTCCTACGGCATTITTTCT	ChIP
Fab8 T2	GAATGGAAGCTCTTCGCTTGC	ChIP

**Tabla 1. Lista de oligonucleótidos utilizados.** Cada oligonucleótido se muestra en dirección 5' -> 3' y se indica el experimento para el que se ha utilizado.

## 1.2. Plásmidos

Nombre	Características
pTOPO-cDNAHaspin	Hemos insertado el cDNA de Haspin para secuenciarlo.
pET-Haspin	Hemos insertado la región N-terminal de Haspin para expresar la proteína con un tag de histidinas y purificarla para hacer anticuerpo.
pUASattB-cDNAHaspin	Hemos insertado el cDNA de Haspin tras el promotor UAS para hacer una línea transgénica.
pUASattB-cDNAHaspin-HA	Hemos insertado el cDNA de Haspin junto al tag HA tras el promotor UAS para hacer una línea transgénica.
pattB-PROM-cDNAHaspin	Hemos insertado el promotor y el cDNA de Haspin en un plásmido attB para hacer una línea transgénica y la mutación del dominio kinasa.
pattB-PROM-cDNAHaspin H20A	Hemos insertado el promotor y el cDNA de Haspin con el dominio kinasa mutado en un plásmido attB para hacer una línea transgénica.
pattB-PROM-cDNAHaspin K282M	Hemos insertado el promotor y el cDNA de Haspin con el dominio kinasa mutado en un plásmido attB para hacer una línea transgénica.
pUASattB-cDNAHaspin H20A-HA	Hemos insertado el cDNA de Haspin con el dominio kinasa mutado y el tag HA tras el promotor UAS para hacer una línea transgénica.
pUASattB-cDNAHaspin K282M-HA	Hemos insertado el cDNA de Haspin con el dominio kinasa mutado y el tag HA tras el promotor UAS para hacer una línea transgénica.
cDNA Haspin	<i>Drosophila Genomics Resource Center</i> LD07633
UAS-pds5-HA	<i>Drosophila Genomics Resource Center</i> UFO 12474
UAS-CapH2-HA	<i>Drosophila Genomics Resource Center</i> UFO 09514

**Tabla 2. Lista de plásmidos utilizados.** Se describen las características o el origen de cada plásmido empleado.



### **1.3. Anticuerpos**

Nombre	Características	Origen
CP190	Conejo; Imm 1:500 IP 2µl WB 1:5000	Generado en el laboratorio
Ibf1	Rata; Imm 1:400 IP 2µl	Generado en el laboratorio
Ibf2	Rata; Imm 1:400 IP 2µl	Generado en el laboratorio
Pita	Conejo; IP 2µl WB 1:2000	Glover D.M.
Haspin	Rata	Generado en el laboratorio
H3T3ph	Conejo; Imm 1:100	Upstate 07-424
	Conejo; IP 3µl	Millipore 04-746
CenpC	Rata; Imm 1:200	Cedido por Olga Moreno-Moreno y Mònica Torras-Llort
HA	Rata; Imm 1:100 WB: 1:500	Roche (3F10)
H3	Conejo; WB 1:2500	Cell signaling 9715
Lamina	Ratón; Imm 1:500 WB 1:5000	DSHB ADL67.10-s
GFP	Ratón; Imm 1:50	ROCHE 11 814 460 001
	Conejo; Imm 1:100	Invitrogen AIII22
Pc	Conejo; IP 3µl	Santa Cruz Biotechnologysc-25762
BEAF	Conejo; Imm 1:100	DSHB BEAF
actina	Conejo; WB: 1:1000	Sigma 2066

**Tabla 3. Lista de anticuerpos primarios utilizados.** Se indica el animal, título y técnica en la que ha sido empleado. En la última columna se describe su origen. Imm: inmunolocalización. IP: inmunoprecipitación. WB: Western Blot.

Animal	Características	Origen
Anti-conejo HRP	WB 1:10000	Jackson (111-035-144)
Anti-rata HRP	WB 1:10000	Jackson (712-035-150)
Anti-ratón HRP	WB 1:10000	Jackson (715-035-150)
Anti-conejo Cy2	Imm 1:400	Jackson (111-225-144)
Anti-conejo Cy3	Imm 1:400	Jackson (111-165-144)
Anti-rata Cy2	Imm 1:400	Jackson (712-225-150)
Anti-rata Cy3	Imm 1:400	Jackson (712-165-150)
Anti-ratón Cy2	Imm 1:400	Jackson (115-225-146)
Anti-ratón Cy3	Imm 1:400	Jackson (115-165-146)

**Tabla 4. Lista de anticuerpos secundarios utilizados.** Se indica la técnica, título y origen. WB: Western Blot. Imm: inmunolocalización.

#### 1.4. Líneas de *Drosophila*

Nombre	Descripción	Origen
<i>w<sup>-</sup> (w<sup>1118</sup>)</i>	Mutación de <i>white</i>	<i>Bloomington Drosophila Stock Center</i>
<i>yw</i> ; <i>CyoPBactransposasa/Sp</i>	Contiene una transposasa capaz de hacer escindir un elemento transponible	<i>Bloomington Drosophila Stock Center #8285</i>
<i>H<sub>2</sub>Av-GFP</i>	Expresa GFP unido a <i>H<sub>2</sub>Av</i>	Clarkson and Saint, 1999
<i>UAS-GFP-RNAi</i>	RNAi contra GFP bajo control de la secuencia UAS	<i>Bloomington Drosophila Stock Center</i>
<i>hsp70-Gal4</i>	Expresa la proteína Gal4 bajo <i>heat-shock</i>	<i>Bloomington Drosophila Stock Center #2077</i>
<i>phm-Gal4</i>	Expresa Gal4 en glándula protorácica bajo control del promotor de <i>phm</i>	Rewitz et al. 2009
<i>lio-Gal4</i>	Expresa Gal 4 en disco imaginal y glándula salivar bajo control del promotor de <i>lio</i>	<i>Bloomington Drosophila Stock Center #4669</i>
<i>da-Gal4</i>	Expresa Gal4 de manera ubicua bajo control del promotor de <i>da</i>	<i>Bloomington Drosophila Stock Center #8641</i>
<i>arm-Gal4</i>	Expresa Gal4 de manera ubicua bajo control del promotor de <i>arm</i>	<i>Bloomington Drosophila Stock Center</i>
<i>nub-Gal4; UAS- Dcr-2</i>	Expresa Gal4 en disco imaginal y glándula salivar bajo control del promotor de <i>nub</i> , sobreexpresando dicer 2	Cedido por Andreu Casali
<i>69B-Gal4</i>	Expresa Gal4 de manera ubicua bajo control del promotor de <i>69B</i>	<i>Bloomington Drosophila Stock Center</i>
<i>act5C-Gal4</i>	Expresa Gal4 de manera ubicua bajo control del promotor de <i>act</i>	<i>Bloomington Drosophila Stock Center</i>
<i>B7; da-Gal4</i>	Expresa Gal4 de manera ubicua bajo control del promotor de <i>da</i> , con Fab7 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio
<i>UAS-CTCF</i>	Expresa <i>CTCF</i> bajo control de la secuencia UAS	Mohan et al., 2007
<i>UAS-CTCF-RNAi</i>	RNAi contra <i>CTCF</i> bajo control de la secuencia UAS	Vienna Drosophila RNAi Center-VDRC #30713
<i>CTCF<sup>y+6</sup></i>	Mutante de <i>CTCF</i>	Gerasimova et al., 2007
<i>UAS-CP190-RNAi</i>	RNAi contra <i>CP190</i> bajo control de la secuencia UAS	Vienna Drosophila RNAi Center-VDRC #35078
<i>CP190<sup>H31-2</sup></i>	Mutante de <i>CP190</i>	Pai et al., 2004
<i>CP190<sup>P1</sup></i>	Mutante de <i>CP190</i>	Pai et al., 2004
<i>CP190<sup>P11</sup></i>	Mutante de <i>CP190</i>	Pai et al., 2004

## Materiales y métodos

<i>UAS-Ibf1-RNAi</i>	RNAi contra <i>Ibf1</i> bajo control de la secuencia UAS	<i>Vienna Drosophila RNAi Center</i> -VDRC #35426
<i>UAS-Ibf2-RNAi</i>	RNAi contra <i>Ibf2</i> bajo control de la secuencia UAS	<i>Vienna Drosophila RNAi Center</i> -VDRC #42121
<i>Ibf1</i> <sup>ΔRB1</sup>	Mutante de <i>Ibf1</i>	Cuartero et al. 2014
<i>Ibf2</i> <sup>GSV17</sup>	Mutante de <i>Ibf2</i>	Cuartero et al. 2014
<i>Ibf2</i> <sup>GSV3</sup>	Mutante de <i>Ibf2</i>	Cuartero et al. 2014
<i>pita</i> <sup>2</sup>	Mutante de <i>pita</i>	Page et al. 2005
<i>pita</i> <sup>3</sup>	Mutante de <i>pita</i>	Page et al. 2005
<i>UAS-Haspin-RNAi</i>	RNAi contra <i>Haspin</i> bajo control de la secuencia UAS	<i>Bloomington Drosophila Stock Center</i> #35276
<i>Haspin</i> <sup>PROM</sup>	Expresa <i>Haspin</i> bajo control de su propio promotor	Creado en el laboratorio
<i>UAS-Haspin-HA</i>	Expresa <i>Haspin</i> unido al tag HA bajo control de la secuencia UAS	Creado en el laboratorio
<i>Haspin</i> <sup>86</sup>	Mutante de <i>Haspin</i>	Creado en el laboratorio
<i>Haspin</i> <sup>128</sup>	Mutante de <i>Haspin</i>	Creado en el laboratorio
<i>Haspin</i> <sup>H420A</sup>	Expresa <i>Haspin</i> bajo control de su propio promotor con el dominio kinasa mutado	Creado en el laboratorio
<i>Haspin</i> <sup>K282M</sup>	Expresa <i>Haspin</i> bajo control de su propio promotor con el dominio kinasa mutado	Creado en el laboratorio
<i>UAS-Haspin-HA</i> <sup>H420A</sup>	Expresa <i>Haspin</i> unido al tag HA con el dominio kinasa mutado bajo control de la secuencia UAS	Creado en el laboratorio
<i>UAS-Haspin-HA</i> <sup>K282M</sup>	Expresa <i>Haspin</i> unido al tag HA con el dominio kinasa mutado bajo control de la secuencia UAS	Creado en el laboratorio
<i>pds5</i> <sup>e3</sup>	Mutante de <i>pds5</i>	<i>Bloomington Drosophila Stock Center</i> #26162
<i>UAS-pds5-RNAi</i>	RNAi contra <i>pds5</i> bajo control de la secuencia UAS	<i>Bloomington Drosophila Stock Center</i> #35362
<i>UAS-pds5-HA</i>	Expresa <i>pds5</i> unido al tag HA bajo control de la secuencia UAS	Creado en el laboratorio a partir de plásmido comercial
<i>SMC1</i> <sup>exc46</sup>	Mutante de <i>SMC1</i>	<i>Bloomington Drosophila Stock Center</i> #25718
<i>UAS-CapH2-RNAi</i>	RNAi contra <i>CapH2</i> bajo control de la secuencia UAS	<i>Bloomington Drosophila Stock Center</i> #38196
<i>UAS-CapH2-HA</i>	Expresa <i>CapH2</i> unido al tag HA bajo control de la secuencia UAS	Creado en el laboratorio a partir de plásmido comercial
F6 133.2	Contiene Fab6 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio

F6 26.3C	Contiene Fab6 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio
F7 B+P 2.5	Contiene Fab7 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio
F7 27.1 B7	Contiene <i>insulator</i> del Fab7 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio
F8 134.1	Contiene Fab8 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio
F8 59.2	Contiene Fab8 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio
F8 109.6	Contiene Fab8 entre el <i>enhancer</i> del gen <i>white</i> y el gen	Creado en el laboratorio

Tabla 5. Líneas de *Drosophila melanogaster* utilizadas. Se indica el nombre, características y origen.

### 1.5. Cepas bacterianas

- Para amplificar plásmidos hemos utilizado **DH5 $\alpha$** : *deoR endA1 gyrA96 hsdR17* ( $r^{-}K m^{-}K$ ) *recA1 relA1 supE44 thi-1 F'*[*traD36 proAB+ lac1q lacZ  $\Delta$ M15*].
- Para expresar proteínas hemos utilizado **BL21 (DE3)**: *BF $^{-}$  dcm ompT hsdS* ( $r^{-}B m^{-}B$ ) *gal $\lambda$*  (DE3).

### 1.6. Líneas celulares

En este trabajo hemos empleado siempre la línea S2 de células de *Drosophila*. Son células de origen embrionario tardío (20-24 horas) que crecen en semi-adherencia en los flascos de medio.

## **2. MÉTODOS**

### **2.1. Mantenimiento de cultivos celulares**

Crecemos las células S2 a 25°C sin CO<sub>2</sub> con medio *Schneider's insect Medium* (Sigma) suplementado con 100µl/mL de estreptomicina (Gibco), 100 unidades/mL de penicilina (Gibco) y un 10% de suero fetal bovino (FBS) (Gibco), el cual se ha de inactivar 30 minutos a 56°C. Las mantenemos en flascos Corning de 25cm<sup>2</sup> a una densidad de 1-8X10<sup>6</sup> células/mL en un volumen de 5mL, haciendo diluciones 1/5 cada 3-4 días. Para obtener grandes cantidades de células, preparamos un flascón de 125cm<sup>2</sup> haciendo una dilución 1/5 en un volumen final de 25mL a partir de células bastante confluentes, y las dejamos crecer durante 3-4 días.

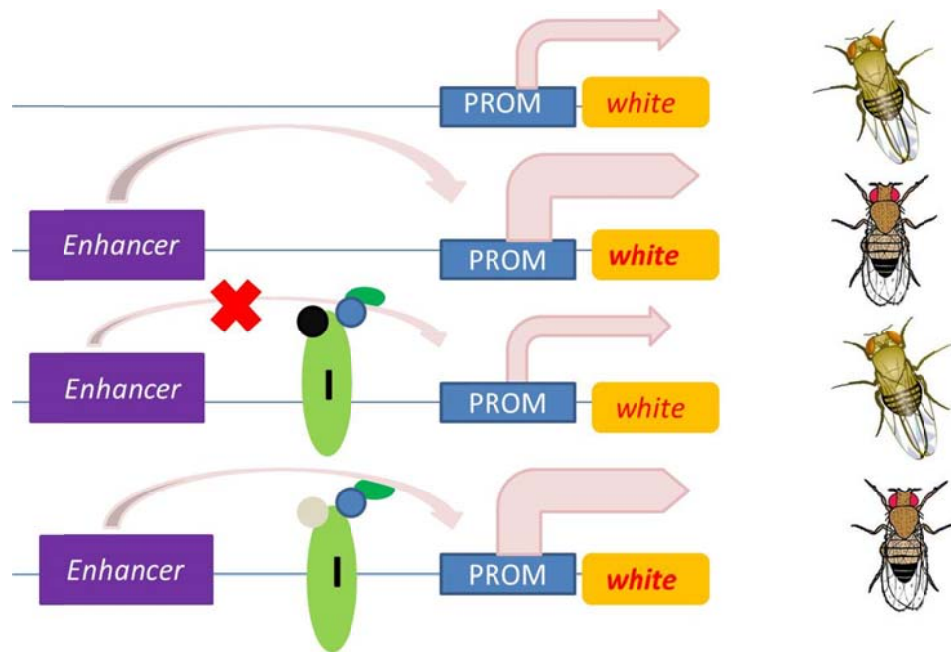
### **2.2. Técnicas de manipulación de *D. melanogaster***

#### **2.2.1. Ensayo *enhancer-blocking***

Mediante esta técnica podemos detectar si una proteína presenta la actividad característica de los *insulators* de interrumpir la comunicación entre un *enhancer* y un promotor (*enhancer-blocking*). Consiste en lo siguiente: la expresión basal de un gen por su promotor (Figura 21, primer panel), se incrementa en presencia de un *enhancer* (Figura 21, segundo panel). Un *insulator* colocado entre el *enhancer* y el promotor impide la comunicación entre estos dos elementos y el gen se expresa a nivel basal (Figura 21, tercer panel). Si una proteína es necesaria para la función de dicho *insulator*, la carencia de la misma hará que no pueda funcionar, por lo que la expresión del gen se verá incrementada gracias a que el *enhancer* sí podrá actuar sobre el promotor (Figura 21, cuarto panel).

Para realizar el ensayo tenemos varias líneas transgénicas en las que hay insertada una construcción con *enhancer* y promotor del gen *reporter white*, y entre ellos un *insulator* diferente en cada una de las líneas (mirar materiales para ver las líneas transgénicas empleadas). Está hecho sobre un fondo mutante para el gen *white* (responsable de la pigmentación de los ojos en *Drosophila melanogaster*), de manera que la coloración solo depende del grado de expresión del gen *reporter*.

Para estudiar la actividad *insulator* de una proteína, cruzamos las líneas transgénicas con moscas control y en paralelo con mutantes de la proteína y comparamos el color de los ojos de los descendientes de los dos cruces. Si la proteína es necesaria para la actividad del *insulator*, éste no funcionará correctamente en las mutantes, por lo que en la descendencia se verá un aumento en la pigmentación de los ojos.



**Figura 21. Ensayo *enhancer-blocking*.** Esquema representativo del fundamento de un ensayo de este tipo. El grosor de las flechas correlaciona con el nivel de expresión del gen *white*, manifestado en la intensidad de la coloración de los ojos de las moscas. I: *insulator*.

### 2.2.2. Cuantificación de la coloración del pigmento de ojos

Recogemos las moscas que provienen del cruce que acabamos de describir, de unas 6 horas, y las dejamos 24h a 18°C para que las diferencias en la coloración sean uniformes. Tras este período, congelamos a -20°C.

Para la cuantificación congelamos 20 moscas en nieve carbónica. Sin dejar que se descongelen, vortexamos a máxima intensidad hasta que las cabezas se separen de los cuerpos. Las ponemos en un eppendorf y centrifugamos al máximo 5 minutos. Añadimos 15µl de etanol 30% a pH=2, trituramos y vortexamos a muy baja intensidad, centrifugando después durante 20 minutos también a baja intensidad. Añadimos 0.3µl de H<sub>2</sub>O<sub>2</sub> al 0.5% para que todo el pigmento esté en el mismo estado de oxidación. Vortexamos suavemente y dejamos las muestras 24 horas a 25°C envueltas en papel de aluminio para que el pigmento quede protegido de la luz. A las 24 horas centrifugamos al máximo durante 30 minutos. Para cada muestra hacemos 3 lecturas del sobrenadante, absorbancia a  $\lambda=480$  nm en nanodrop 1000 3.7 aplicación UV-Vis.

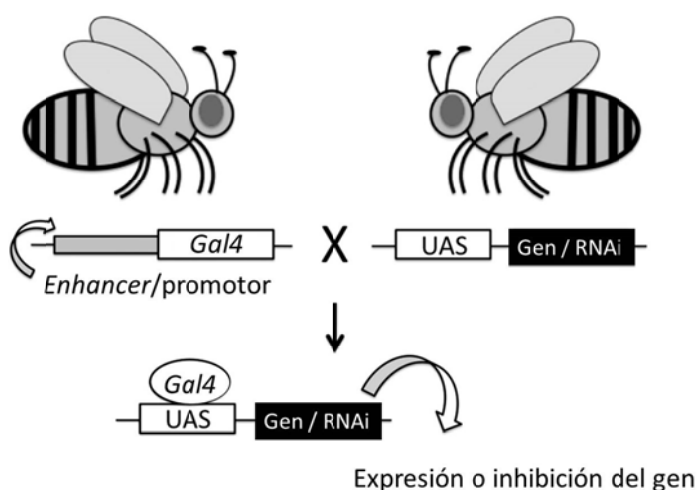
### 2.2.3. Sistemas de control de la expresión de genes en *Drosophila*

#### 2.2.3.1. Heat-shock

Este método permite la expresión de un transgén bajo un promotor que se activa mediante un choque térmico (*heat-shock*), de manera que se puede controlar el momento, así como el nivel y persistencia de su expresión variando la intensidad y la duración de este choque de temperatura.

#### 2.2.3.2. Sistema UAS-Gal4

Es una técnica que permite tanto la expresión ectópica de genes así como la degradación del RNA mensajero en una amplia variedad de tejidos. Mediante un cruce genético entre una mosca que contiene el gen *Gal4* unido a un promotor y otra en cuyo genoma está insertado el promotor UAS (*upstream activating sequence*) seguido de la secuencia del gen que queremos expresar o el RNAi del que queremos que no se exprese la proteína, ponemos en contacto UAS y *Gal4* (Figura 22). De esta manera, la secuencia que se encuentra tras UAS solo será expresada en los tejidos en los que esté presente *Gal4*. El promotor de *Gal4* se puede expresar en todo el cuerpo o bien ser específico de un tejido en concreto, controlando la localización del efecto que estamos induciendo de manera específica.



**Figura 22. Sistema UAS-Gal4 para expresar o inhibir un gen.** Esquema representativo del fundamento de un ensayo de este tipo.

#### **2.2.4. Estudio de viabilidad**

Para ver si la carencia de una proteína afecta a la viabilidad de las moscas, llevamos a cabo cruces entre moscas con el mismo genotipo, que lleven la mutación en estudio balanceada. Calculamos el porcentaje teórico que obtendríamos de los descendientes de cada uno de los posibles genotipos de ser 100% viable. Comparando el valor obtenido con el previsto, podemos conocer si la proteína afecta a la viabilidad o no.

#### **2.2.5. Duración de la vida de moscas adultas**

Para analizar el efecto de la carencia de una proteína sobre la supervivencia de una mosca adulta, cogemos un número representativo de moscas control y mutantes de la misma edad, y de manera rutinaria vamos tomando nota del momento en que mueren para ver si hay alguna diferencia debido a la carencia de la proteína.

#### **2.2.6. Estudio del tiempo de desarrollo**

Para realizar los experimentos de retraso en el desarrollo con larvas de exactamente la misma edad, hacemos puestas en placas de agar con zumo de manzana durante dos horas. Las larvas de primer estadio, que están sincronizadas, las colocamos en medio estándar con levadura a 25°C.

Para los análisis del tiempo de desarrollo, ponemos 30 larvas por tubo y evaluamos el paso de larva en estadio II a III así como la pupación, de manera periódica.

Para comprobar cambios en el desarrollo producidos por la alimentación procedemos de la misma manera, y la alimentación la hacemos añadiendo 20-hidroxiecdisona a una concentración de 0.33mg/ml en etanol y/o colesterol a 0.35mg/ml. Las larvas control las alimentamos con etanol.

### **2.3. Técnicas de trabajo con DNA**

#### **2.3.1. Reacción en cadena de la polimerasa (PCR)**

Las PCR clásicas las hacemos con unas condiciones estándar, a las que realizamos algunas variaciones en función de los oligonucleótidos empleados (aumentando o disminuyendo la temperatura de hibridación, y en algún caso variando ligeramente la cantidad de magnesio) o cambios según la largura del amplicón (aproximadamente 1kb por minuto de elongación).



## **Materiales y métodos**

En una reacción estándar ponemos el DNA (del orden de nanogramos),  $MgCl_2$  a 2mM, dNTP a 0.4mM, el tampón a 1X, los oligonucleótidos a 0.5pmol/ $\mu$ l, y la polimerasa a 0.05U/ $\mu$ l. Generalmente hemos usado la *Phusion* DNA polimerasa.

Los ciclos y temperaturas estándar que utilizamos:

2 minutos a 95°C

1 min a 95°C... Desnaturalización	} 30 ciclos
1 min a 55°C... Hibridación	
1 min a 72°C... Elongación	

10 min a 72°C

### **2.3.2. Transformación de células competentes**

Incubamos las células competentes con 10 $\mu$ l del producto de ligación durante 30 minutos en hielo. Realizamos un choque térmico en un baño a 42°C durante 2 minutos y, seguidamente, añadimos 400 $\mu$ l de medio LB sin ningún tipo de antibiótico y lo incubamos durante 30 minutos a 37°C. Finalmente extendemos las células en placas de LB con el antibiótico al que es resistente nuestro plásmido y las incubamos durante toda la noche en una cámara a 37°C.

#### Medio LB:

Triptona 10g/L

Extracto de levadura 5g/L

NaCl 10g/L

NaOH 1mM

### **2.3.3. Minipreparación de DNA plasmídico**

Inoculamos 4mL de medio LB con antibiótico con una colonia aislada de la placa de transformación. Dejamos crecer toda la noche a 37°C y en agitación. A la mañana siguiente pasamos 1.5mL a un eppendorf y centrifugamos a 9000rpm durante un minuto. Quitamos el sobrenadante y resuspendemos en 100 $\mu$ l de GTE vortexando. Añadimos 200 $\mu$ l de solución de NaOH 0.2N-SDS 1% y lo dejamos en hielo durante cinco minutos. Adicionamos 150 $\mu$ l de acetato de sodio 3M, agitamos y dejamos en hielo cinco minutos más. Centrifugamos cinco minutos a 14000rpm y pasamos el sobrenadante a un eppendorf nuevo. Añadimos 450 $\mu$ l de fenol/cloroformo, vortexamos y centrifugamos 2 minutos a 14000rpm. Recuperamos la fase acuosa (superior) y le añadimos 1mL de etanol frío. Invertimos dos o tres veces y centrifugamos 5 minutos a 14000rpm. Quitamos el sobrenadante y lavamos el pélet con 300 $\mu$ l de etanol 70%. Centrifugamos 3 minutos a 14000rpm,

descartamos el sobrenadante y dejamos secar el pélet. Lo resuspendemos con 50µl de TE 1X y 1µg de RNAsa, dejándola actuar durante 15 minutos a 37°C.

### TE 1X:

10mM Tris HCl pH 8.0

1mM EDTA

### GTE:

50mM Glucosa

10mM EDTA

0.25mM Tris HCl pH 8.0

### **2.3.4. Maxipreparación de DNA plasmídico**

Inoculamos 500mL de medio LB con antibiótico con un minicultivo. Incubamos toda la noche a 37°C en agitación. Centrifugamos a 4500rpm. Descartamos el medio, extraemos y purificamos el DNA con *Nucleobond*. Dejamos secar el pélet de DNA y lo resuspendemos en 500µL de TE 1X.

### **2.3.5. Extracción de DNA genómico de mosca**

Ponemos una mosca adulta en un eppendorf y la congelamos a -20°C. Añadimos 40µL de solución SB y 10µL de proteinasa K 20mg/mL y después la disgregamos utilizando un émbolo para eppendorfs. Incubamos la mezcla resultante a 37°C durante 30 minutos y, justo después, a 95°C durante 3 minutos para inactivar la proteinasa K. A partir del material resultante realizamos una dilución 1/5 para utilizar como DNA sobre el que hacer experimentos de PCR.

### Solución SB:

10mM Tris-HCl pH 8.0

1mM EDTA

25mM NaCl

### **2.3.6. Extracción de DNA genómico de glándulas salivares**

Disecionamos 20 glándulas salivares y las ponemos en 300µl de tampón de disección en hielo. Tras 30 minutos, añadimos 100µl de stop mix. Lo mezclamos durante 5 minutos a temperatura ambiente. Añadimos 20µg de proteinasa K y digerimos durante 2 horas a 45°C. Extraemos el DNA mediante fenol/cloroformo. Precipitamos con etanol y acetato de sodio, y tras limpiar el DNA resuspendemos en 20µl de TE y cuantificamos la cantidad de DNA extraída por PCR cuantitativa.

## **Materiales y métodos**

### Tampón de disección:

NaCl 130mM

KCl 5mM

CaCl<sub>2</sub> 1.5mM

### 1ml de stop mix:

40μL 1M Tris pH 7.5

8μL 5M NaCl

160μL 0.5M NaEDTA

200μL 10% SDS

56μL β-mercaptoetanol

536μL H<sub>2</sub>O

### **2.3.7. Cuantificación del tamaño de los núcleos de glándulas salivares**

Preparamos muestras de glándulas salivares (apartado 2.4.5.3.). Tomamos las fotos en el microscopio confocal Leica TCS SP2-AOBS con el objetivo de 63X haciendo siempre el mismo grosor de cortes. Una vez tenemos las fotos, sumamos el mismo número de cortes entre muestras que queremos comparar y delimitamos el área de los núcleos marcados con DAPI. Calculamos el área de esta señal con el programa Fiji. Para representar comparaciones entre diferentes muestras, utilizamos el programa GenePattern para realizar una gráfica *Box Plot*.

## **2.4. Técnicas de trabajo con RNA**

### **2.4.1. Extracción de RNA**

En caso de hacer la extracción de larvas enteras o glándulas protorácicas, la realizamos a partir de 10 larvas en estadio III o de cerebros con las glándulas. Si la hacemos de glándulas salivares, empleamos 15 pares de glándulas.

Para purificar, añadimos 0.5mL de trizol y lo pasamos 20 veces por una jeringa de 1mL con una aguja de 0.8mm de diámetro para larvas enteras, y de 0.5mm para cerebros y glándulas salivares. Lo dejamos 5 minutos a temperatura ambiente, añadimos 0.1mL de cloroformo y agitamos con la mano 2 minutos a temperatura ambiente. Centrifugamos al máximo 15 minutos a 4°C. Recuperamos la fase acuosa y añadimos un volumen de solución RLT y 2% de β-mercaptoetanol, continuando la purificación con el kit *RNAeasy miniKit* de *Qiagen*. Al finalizar, calculamos la cantidad extraída en nanodrop 1000 3.7 aplicación UV-Vis, absorbancia a λ=260nm.

### 2.4.2. Retrotranscripción y PCR cuantitativa

Se trata, en primer lugar, de retrotranscribir el RNA a cDNA y, en segundo lugar, amplificar este DNA para poder estimar las cantidades relativas de diferentes moléculas de RNA. La PCR cuantitativa funciona de manera que hay un reactivo, el *SybrGreen*, que se une al DNA de doble cadena, en este caso se une a los amplicones resultantes de la PCR, y después, emite fluorescencia. Ésta es detectada por el mismo termociclador que nos informa de la cantidad del amplicón generado a través de los ciclos de PCR. La estimación de las cantidades relativas de DNA inicial surge de la idea que, dado un número de ciclos igual para diferentes muestras, aquellas que den más fluorescencia partían de una cantidad mayor de DNA inicial, y viceversa. Por tanto, no se trata de una estimación de la cantidad de DNA de la que se originaba cada muestra sino de una estimación relativa entre muestras de un mismo experimento. Es por eso que a menudo utilizamos genes de transcripción constitutiva como gen de referencia, ya que los resultados de un experimento se dan como una medida de la relación de los genes de interés a un gen de referencia.

La retrotranscripción se hace con el *kit Transcriptor First Strand cDNA Synthesis* de Roche. Utilizamos 1µg de RNA, oligo-dT (que se une a todos los RNA poli-A) o bien oligo-random (que no se une a las poli-A sino de manera aleatoria en diferentes zonas del RNA), 2.5µM, tampón de transcripción a 1X, inhibidor de RNAsa 20U, una mezcla de dNTP 1mM y 10U de la enzima transcriptasa reversa. Aumentamos el volumen con H<sub>2</sub>O hasta 20µL y dejamos la reacción durante 1 hora a 50°C. Después, inactivamos la enzima calentando las muestras a 85°C durante 5 minutos y dejándolo en hielo.

Para hacer la PCR utilizamos el *kit Light Cyclor 480 Sybr Green I Master*. Preparamos las reacciones en un volumen de 10µL: 2µL de cDNA, 0.5µL de cada primer, 5µL de *SybrGreen* y 2µL de H<sub>2</sub>O. Realizamos las reacciones de PCR con 2 réplicas técnicas y varias réplicas biológicas. Ponemos las reacciones en placas de 12X8 pocillos y utilizamos el termociclador *Light Cyclor 480* de Roche. Analizamos los datos de expresión mediante el método de las  $\Delta\Delta Ct$ .

El programa estándar es el siguiente:

5 minutos a 95°C

10 segundos a 95°C... Desnaturalización	}	X 45 ciclos
10 segundos a 60°C... Hibridación		
10 segundos a 72°C... Elongación		

## **Materiales y métodos**

### **2.4. Técnicas de trabajo con proteínas**

#### **2.4.1. Extracción proteica**

Si extraemos proteínas de cerebro de larva, diseccionamos 20 larvas en estadio III y aislamos los cerebros en PBS 1X. Si lo hacemos de glándulas salivares, necesitamos 25 larvas. Ponemos estos tejidos en 50µL de PLB 1X-β-mercaptoetanol 2%. Sonicamos las muestras en un sonicador Branson mediante un pulso a una amplitud del 30%. Lo dejamos durante 3 minutos a 95°C y centrifugamos 1 minuto a 9000rpm.

##### PLB 5X

0.125M Tris pH 6.8

22% glicerol

5% SDS

Azul de bromofenol (punta de espátula)

#### **2.4.2. Electroforesis en gel de acrilamida**

Esta técnica separa proteínas según su peso molecular. La separación tiene lugar en condiciones desnaturizantes con SDS y a través de una matriz de poliacrilamida (PAGE). El gel está compuesto de una matriz apiladora, al 4% de acrilamida, y una separadora al 10% (en función del peso de la proteína que nos interesa detectar se puede hacer de menor concentración (para proteínas de mayor peso) o de mayor (para proteínas de bajo peso). Ponemos todas las muestras en tampón de carga (PLB) con β-mercaptoetanol. Tras haber estado 5 minutos a 95°C, las corremos a través del gel con tampón Laemmli 1X.

##### Gel apilador:

Acrilamida:bisacrilamida 29:1 4%

0.125M Tris-HCl pH 6.8

0.1% SDS

6µL PSA 10%/mL solución

3µL TEMED/mL solución

##### Gel separador:

Acrilamida:bisacrilamida 29:1 10%

0.4M Tris-HCl pH 8.7

0.1% SDS

6µL PSA 10%/mL solución

3µL TEMED/mL solución

Tampón Laemmli:

1% SDS  
2M Glicina  
250mM Tris base

**2.4.3. Western Blot**

Una vez hemos separado las proteínas con un gel de acrilamida (SDS-PAGE), las transferimos del gel a una membrana de nitrocelulosa durante 1 hora a 4°C a 100V en tampón de transferencia. Después bloqueamos la membrana con PBS-*tween* 0.1%-leche 10% durante al menos una hora a temperatura ambiente. A continuación incubamos con el anticuerpo primario diluido en PBS-*tween* 0.1%-leche 5% durante toda la noche a 4°C. Se hacen tres lavados de 10 minutos con PBS-*tween* 0.1% e incubamos durante una hora con el anticuerpo secundario conjugado con peroxidasa. Lavamos 5 veces de cinco minutos con PBS-*tween* 0.1%. Para revelar añadimos sobre la membrana unas gotas de la solución reveladora Inmobilon Western (Millipore) de manera que recubrimos toda la membrana, durante un minuto. Quitamos el exceso y lo exponemos a un film de autoradiografía. El film se revela con un *HyperProcessor de Amersham*.

Tampón de transferencia:

25mM Tris Base  
192mM glicina  
15% metanol

Solución de bloqueo:

10% leche en polvo  
0.05% de *tween*20 en PBS 1X

**2.4.4. Obtención y purificación de anticuerpos****2.4.4.1. Expresión y purificación de proteínas fusionadas a una cola de histidinas**

La expresión de proteínas clonadas en los vectores pET se encuentra bajo el control de señales de transcripción y traducción del bacteriófago T7. La cepa de *Escherichia coli* empleada para expresar, BL21, contiene una copia cromosomal del gen de la T7 RNA polimerasa bajo el control del promotor *lacUV5*, el cuál es inducible por IPTG. Gracias a la elevada actividad de la RNAPol T7, después de unas pocas horas de inducción la proteína que queremos expresar puede constituir más del 50% del contenido proteico total bacteriano.

Expresamos el extremo amino-terminal de Haspin (desde el aminoácido 18 que codifica la metionina para el inicio de la transcripción hasta el de la posición 334) fusionado a una cola de

## **Materiales y métodos**

histidinas por el lado carboxi-terminal con el objetivo de purificarlo y utilizarlo para la obtención de anticuerpos policlonales específicos para Haspin. Lo hacemos contra esta parte de la proteína con el fin de evitar el posible reconocimiento de otras proteínas con actividad kinasa por la homología que presentan entre ellas en su dominio catalítico, en el extremo C-terminal. Por ello, contruimos el plásmido de expresión pET-Haspin y expresamos la proteína de fusión recombinante en la cepa BL21.

### ***Pruebas de expresión***

Primero hacemos una prueba de expresión a pequeña escala para encontrar las condiciones de inducción más adecuadas. Entonces, tras transformar la cepa BL21 con el plásmido pET-Haspin inoculamos 3 cultivos de 4mL de LB/kanamicina 50µg/mL /cloramfenicol 25µg/mL con 3 colonias aisladas y los dejamos crecer a 37°C durante toda la noche. Al día siguiente usamos 200µL para reinocular minicultivos nuevos, los dejamos crecer a 37°C hasta que alcanzan una DO<sub>600</sub> de 0.6-0.8. Entonces inducimos la expresión con IPTG 1mM. Cogemos alícuotas de 400µL justo antes de inducir el cultivo, y 1, 2 y 3 horas después. Las centrifugamos durante 1 minuto a 9000rpm, las resuspendemos en 100µL de tampón de carga PLB 5X y un 10% de β-mercaptoetanol y, finalmente, sonicamos durante 5 segundos. A continuación analizamos 25µL de cada una de las alícuotas por SDS-PAGE (ver apartado 2.4.2.).

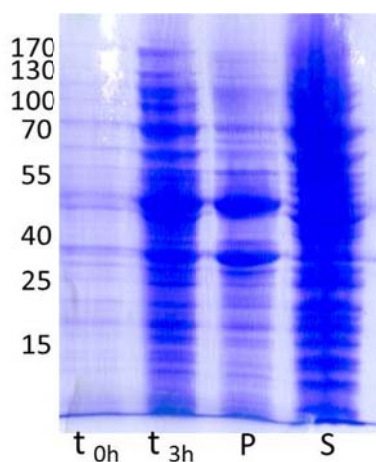
### ***Expresión a gran escala***

Una vez determinadas las condiciones de expresión procedemos a la producción de grandes cantidades del extremo N-terminal de Haspin fusionado con una cola de histidinas. Volvemos a transformar en la cepa BL21 con el plásmido pET-Haspin e inoculamos un cultivo de 4mL de LB/kanamicina 50µg/mL / cloramfenicol 25µg/mL con 3 colonias aisladas y lo dejamos crecer a 37°C durante toda la noche. Al día siguiente lo añadimos a 500mL de LB y lo dejamos crecer a 37°C hasta una DO<sub>600</sub> de 0.6-0.8. Entonces, cogemos una alícuota de 400µL, inducimos la expresión con IPTG 1mM y lo dejamos a 37°C durante 3 horas más. Cogemos otra alícuota de 400µL y recogemos las células centrifugando durante 10 minutos a 4000rpm en el rotor JA-14 (Beckman) a 4°C. Resuspendemos el pélet celular en 100mL de STE y volvemos a centrifugar. Congelamos el sedimento celular obtenido en nieve carbónica y lo guardamos a -80°C durante toda la noche para facilitar la lisis.

Al día siguiente lo resuspendemos con 10L de tampón de lisis y lo lisamos sonicando en hielo hasta que la muestra adquiere una consistencia más líquida y un color más oscuro. Empleamos el sonicador Digital Sonifier® (Branson) y sonicamos con el tip grueso (12mm de diámetro) haciendo 10 ciclos de 10 segundos cada uno con intervalos de descanso de 30 segundos y un valor de salida entre 50 y 60 W. Después centrifugamos a 35000rpm con el rotor SW41 (Beckman) durante 1 hora a 4°C

para separar las fases soluble e insoluble (sobrenadante y pélet, respectivamente). Finalmente, congelamos la muestra con nieve carbónica y la guardamos a -80°C.

Para analizar si la proteína de interés se ha expresado de forma soluble o formando cuerpos de inclusión cargamos en SDS-PAGE un poco de pélet (cogemos un poco con una punta amarilla, lo resuspendemos en 400µl de tampón de carga y analizamos 10µL de una dilución 1/10), 10µL de la fracción soluble así como 25µl de las muestras de antes y después de la inducción (Figura 23). Teñimos el gel con una solución fresca de Coomassie 0.05% en agua durante unos diez minutos.



**Figura 23. El extremo N-terminal de Haspin se expresa de forma insoluble.** Gel SDS-PAGE que muestra la expresión de la proteína antes y tras 3 horas de inducción, así como su presencia en el pélet (P) y en la fracción soluble (S).

### STE:

NaCl 0.1M

Tris 10mM pH8

EDTA 1mM

### Tampón de lisis:

NaCl 0.3M

Glicerol 20%

Hepes/NaOH 20mM

EDTA 1mM

β-mercaptoetanol 15mM

PMSF 1mM

NP40 0.1%



## **Materiales y métodos**

### ***Purificación a partir de la fracción insoluble***

La utilización de una cola de histidinas como tag permite la purificación de la proteína de fusión mediante cromatografía de afinidad con una columna de  $\text{Ni}^{2+}$ , el cuál interacciona con los grupos imidazoles de las histidinas haciendo que éstas queden retenidas a la columna.

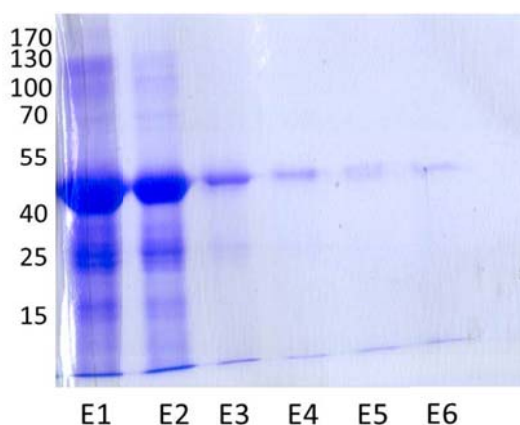
Resuspendemos el pélet en 10mL de tampón de cloruro de guanidinio en agitación durante toda la noche a temperatura ambiente. A la mañana siguiente, centrifugamos durante 1 hora a temperatura ambiente a 11000rpm.

En nuestro caso hacemos la purificación a temperatura ambiente utilizando 1mL de resina Ni-NTA Agarose (Qiagen), que empaquetamos en una columna *Econocolumn* (Biorad) y equilibramos con 10mL de tampón de cloruro de guanidinio. Este paso y los siguientes los realizamos con la ayuda de una bomba peristáltica LKB Pump P-1 (Pharmacia) a velocidad entre 6 y 8 de la posición 1.

Una vez empaquetada la resina, pasamos hasta 3 veces los 10mL de la muestra correspondiente a la fracción insoluble y recogemos la fracción *Flow Through* (FT). Posteriormente hacemos dos lavados: el primero con 10mL de tampón de cloruro de guanidinio (W1) y el segundo con 5mL de tampón urea (W2).

Finalmente eluimos la proteína por competición con Imidazol en un proceso de 3 pasos. Primero hacemos una pre-elución con 3mL de tampón urea con 20mM de Imidazol, recogiendo las fracciones de 1mL (PE1, PE2, PE3). Este paso con baja concentración de Imidazol sirve para eliminar contaminantes que se unen a la resina con una afinidad menor que la cola de histidinas. Después eluimos con 5mL de tampón urea con 100mM de Imidazol (E1-E5), seguida de otra elución con 5mL de tampón urea con 500mM Imidazol (E6-E10).

A continuación, para comprobar el funcionamiento de la cromatografía analizamos 10 $\mu\text{L}$  de cada una de las fracciones por SDS-PAGE. Teñimos el gel con una solución fresca de Coomassie 0.05% en agua durante unos diez minutos (Figura 24 muestra las eluciones desde E1 a E6). Congelamos el resto de las muestras con nieve carbónica y las guardamos a  $-80^\circ\text{C}$ .



**Figura 24. Purificación de la proteína de fusión Haspin-His.** Gel SDS-PAGE donde se muestra las distintas eluciones de la proteína (E1-E6).

### Cloruro de guanidinio:

Cloruro de guanidinio 6M  
 $\beta$ -mercaptoetanol 10mM  
Hepes pH 7.9 20mM  
NaCl 100mM

### Tampón urea:

Urea 8M  
KCl 100mM  
Hepes pH 7.9 20mM  
 $\beta$ -mercaptoetanol 10mM

#### 2.4.4.2. Obtención de anticuerpos policlonales

Obtenemos anticuerpos policlonales específicos contra el extremo N-terminal de Haspin inyectando la proteína recombinante en ratas. Para ello, expresamos la proteína de fusión en *Escherichia coli* como acabamos de redactar, y la cargamos en un gel SDS-PAGE del 10% de acrilamida, 20X20cm y un grossor de 1.5mm, usando como pozo un peine de los *MiniProtean* puesto del revés. Teñimos el gel con una solución fresca de Coomassiae 0.05% en agua durante unos diez minutos y recortamos la banda de interés. Entonces, la rompemos en trozos pequeños y la pasamos por una jeringa de 2mL, depositándola en una eppendorf de 1.5mL. Para que la acrilamida quede lo más fina posible evitando posibles obturaciones de la aguja a la hora de inmunizar la rata, hacemos un agujero con una aguja en la base de este tubo, lo colocamos dentro de otro y centrifugamos a 14000rpm durante 2 minutos. Repetimos este proceso dos veces más, disminuyendo cada vez el tamaño del agujero de la base del eppendorf. Finalmente, resuspendemos la muestra en un volumen de PBS 1X haciendo vórtex. Tiene que quedar una solución viscosa que se ha de poder coger con la pipeta. Si no, añadimos más PBS.

Las inmunizaciones en ratas de forma subcutánea las realiza la Dra. Eva Prats, del estabulario del CID-CSIC (Barcelona). Hace una primera inmunización con el antígeno y 1 volumen de adyuvante de Freund completo (Sigma) y, después hace cuatro recordatorios con intervalos de 15 días, estos con adyuvante incompleto (Sigma). El adyuvante de Freund es una emulsión de aceite mineral-agua preparada con aceites no metabolizables que protegen el inmunogén de una degradación acelerada. El adyuvante completo se diferencia del incompleto por la presencia de muramil dipéptido de la bacteria *Mycobacterium tuberculosis*, que hace que se incrementen los niveles de linfoquinas generando una respuesta immune primaria más corta.

Unos 10 días después de la última inmunización realiza sangrados durante tres días consecutivos. Para procesar la sangre extraída de las ratas, la deja reposando durante un par de horas a temperatura ambiente. Posteriormente centrifuga a 2000rpm durante 10 minutos a 4°C y recoge la fase superior que es la correspondiente al suero. Añade azida sódica a una concentración final de 0.02%, alicuota, congela con nieve carbónica y guarda a -80°C.

## **Materiales y métodos**

Finalmente, comprobamos la especificidad del anticuerpo mediante *Western blot* con proteína extraída de cerebros y glándulas salivares de moscas control (*w*) y mutantes de Haspin (*Haspin*<sup>86</sup>) (ver apartado 2.4.1.), sin obtener ninguna banda que parezca específica de la proteína.

### ***Purificación de anticuerpos por afinidad***

Para conseguir anticuerpos específicos contra el extremo N-terminal de Haspin purificamos por afinidad el suero obtenido de la inmunización de las ratas. Consiste en unir un antígeno utilizado en un soporte sólido sobre el cuál se aplica el suero policlonal. De esta manera solo son retenidos en la columna los anticuerpos específicos, los cuales pueden ser eluidos posteriormente.

Concretamente, unimos la proteína de fusión recombinante a la resina Ni-NTA Agarose (Qiagen). Lavamos la columna con 10mL de tampón B y pasamos 2mL de suero. Posteriormente, hacemos otro lavado con 10mL de tampón B y por último otro con 10mL de tampón C. Finalmente, eluimos con una solución de MgCl<sub>2</sub> 4M.

Para comprobar si la purificación ha funcionado hacemos un *Western blot* con la proteína recombinante usando la elución de la columna como anticuerpo primario y comparándolo con el suero. No obtenemos ninguna mejora.

#### Tampón B:

NaCl 150mM

Tris pH7.4 50mM

#### Tampón C:

NaCl 2M

Tris pH7.4 50mM

### **2.4.5. Immunotinciones**

#### **2.4.5.1. Cromosomas politénicos**

Disecionamos las glándulas salivales de larvas en estadio III en tampón Cohen, en el que las dejamos entre 8 y 10 minutos. Hacemos 6 larvas por preparación. Quitamos el tampón y fijamos primero con la solución de fijación 1, durante 2 minutos y después con la solución de fijación 2, durante 3 minutos. Transferimos las glándulas a un cubreobjetos siliconizado y ponemos un portaobjetos por encima. Picamos varias veces con un punzón para abrir los núcleos, giramos el portaobjetos y apretamos fuerte con el dedo. Sumergimos la muestra en nitrógeno líquido y le

quitamos el cubreobjetos, dejándola en un recipiente con PBS-*tween* 0.05%. Hacemos lavados con este tampón y bloqueamos dos veces de 20 minutos con PBS-*tween* 0.05%-BSA 2%. Ponemos el anticuerpo primario diluido en esta solución (cada anticuerpo requiere su propia dilución) con un cubreobjetos por encima y se incuba toda la noche a 4°C en una cámara húmeda. Hacemos 3 lavados de 5 minutos con PBS-*tween* 0.05%, y ponemos el anticuerpo secundario diluido en PBS-*tween* 0.05%, a 1:400, durante 1 hora a temperatura ambiente, también en cámara húmeda. Finalmente lavamos 3 veces con PBS-*tween* 0.05%, una vez con PBS y montamos la preparación con DAPI-Mowiol.

### Tampón Cohen

MgCl<sub>2</sub> 10mM  
Glicerol de sodio 3P 25mM  
CaCl<sub>2</sub> 3mM  
KH<sub>2</sub>PO<sub>4</sub> 10mM  
NP40 0.5%  
KCl 30mM  
Sacarosa 160mM

### Solución de Fijación 1

PBS 10X 25μL  
Formaldehido 37% 50μL  
H<sub>2</sub>O 2.2mL

### Solución de Fijación 2

Ácido acético 1.125mL  
Formaldehido 37% 50μL  
H<sub>2</sub>O 1.325mL

#### 2.4.5.2. *Squash* de cerebros

Disecionamos los cerebros de larvas en estadio III en medio salino 0.07% NaCl. Utilizamos 6 cerebros por preparación. Incubamos con colchemida a 5μg/μl diluida en este medio durante 2 horas, a oscuras, en cámara húmeda y a temperatura ambiente. Hacemos un choque hipotónico con citrato de sodio al 0.05% durante 10 minutos, fijamos los cerebros con formaldehido al 3.7% en medio fresco durante 30 minutos. Transferimos la muestra a un cubreobjetos siliconizado que contiene ácido acético al 60% y hacemos un *squash* como el descrito en politénicos. Finalmente, sumergimos en nitrógeno líquido.

Tras dos lavados de 10 minutos con PBS-Tritón 1%, bloqueamos con PBS-Tritón 1%-leche 1% durante una hora a temperatura ambiente. Después incubamos con el anticuerpo primario a la dilución adecuada hecha en la solución de bloqueo durante toda la noche a 4°C en cámara húmeda.

## **Materiales y métodos**

A la mañana siguiente limpiamos dos veces con PBS-Tritón 1% e incubamos con el anticuerpo secundario a 1:400 durante 1 hora a temperatura ambiente también en cámara húmeda. Lavamos dos veces con PBS-Tritón 1% y otras dos más con PBS. Finalmente montamos en DAPI-Mowiol.

### **2.4.5.3. Discos imaginales de ala y glándulas salivares**

Disecionamos las larvas en estadio III en PBS 1X y las fijamos con paraformaldehído 4%, incubando en agitación durante 20 minutos a temperatura ambiente. Tras 2 lavados con PBS-Tritón 0.3% bloqueamos durante 1 hora a temperatura ambiente con PBS-Tritón 0.3%-BSA 2%. Incubamos con el anticuerpo primario diluido en la solución de bloqueo a la concentración adecuada durante toda la noche a 4°C en agitación. Al día siguiente hacemos dos lavados de 10 minutos cada uno con PBS-Tritón 0.3% e incubamos con el anticuerpo secundario diluido 1:200 en esta solución durante 1 hora a temperatura ambiente y en oscuridad. Lavamos una vez con PBS-Tritón 0.3% y otras 2 con PBS. Incubamos durante 20 minutos con PBS-DAPI. Finalmente montamos las preparaciones con DAPI-Mowiol.

En el caso de no querer hacer inmunotinción, sino solo el marcaje del DNA, procedemos de la misma manera, pero tras los dos primeros lavados con PBS-Tritón 0.3% pasamos directamente a los lavados con PBS y a la incubación con PBS-DAPI, montando finalmente en DAPI-Mowiol.

### **2.4.6. Co-Immunoprecipitación**

Partimos de dos flascos pequeños de células S2 que centrifugamos cinco minutos a 300g y lavamos con PBS 3 veces. Resuspendemos el pélet con tampón de lisis, vortexamos, lisamos con el *dounce* B, y dejamos 30 minutos en la rueda a 4°C para extraer las proteínas. Centrifugamos 10 minutos a 13200rpm a 4°C. A la cantidad de extracto deseada por inmunoprecipitación (IP) le añadimos 3µl de anticuerpo y lo dejamos incubando a 4°C en la rueda durante toda la noche. A la mañana siguiente añadimos 50µl de proteína-A-sefariosa equilibrada con tampón de lisis y dejamos en la rueda a 4°C durante dos horas. Después lavamos las bolas de sefariosa con tampón de lisis tres veces centrifugando entre medio 1 minuto a 3000rpm. Finalmente eluimos añadiendo PLB1X-β-mercaptoetanol 2%, calentando 5 minutos a 95°C y vortexando.

Corremos las eluciones de cada IP en un gel SDS-PAGE y hacemos un *Western Blot* utilizando como anticuerpo primario el de la proteína de interés que queremos saber si se encuentra en la IP de la proteína inmunoprecipitada. Utilizamos como input una cantidad de extracto que corresponde al 10% de la que hemos utilizado para inmunoprecipitar, que también corremos en el SDS-PAGE y la detectamos por *Western Blot*.

### Tampón de lisis:

Tritón 1%  
NaCl 300mM  
Tris pH8 50mM  
EDTA 1mM  
DTT 1mM  
Inhibidor de proteasas 1X

### **2.4.7. Immunoprecipitación de cromatina (ChIP)**

Partimos de  $1 \times 10^8$  células S2 por cada IP que queramos hacer. Añadimos 1.8% de formaldehído para hacer un *crosslinking* durante 10 minutos en agitación a temperatura ambiente, y a continuación paramos la reacción con glicina a 0.125 M. Centrifugamos las células 2 minutos a 1500g a 4°C, y lavamos el pélite con PBS. Resuspendemos en tampón A, dejamos que se limpie durante 10 minutos, y posteriormente con tampón B. Añadimos a las células TE con SDS al 1% y centrifugamos 2 minutos a 1500g a 4°C. A continuación le hacemos dos lavados más solo con TE, para finalmente resuspender en 4ml de TE con PMSF y SDS al 1%. Hacemos dos alícuotas de 2ml y sonicamos con un sonicador bioruptor a máxima intensidad durante 25 minutos a ciclos de 30 segundos ON y 30 segundos OFF. Juntamos los lisados y añadimos Tritón, deoxicolato sódico y cloruro de sodio hasta conseguir un medio con la misma composición que RIPA. Incubamos 10 minutos en la rueda a 4°C, centrifugamos las muestras, y con la cromatina sonicada que se encuentra en el sobrenadante hacemos alícuotas de 0.5mL para inmunoprecipitar, y guardamos un 10% como input.

Incubamos las muestras de cromatina *crosslinkada* con proteína-A-sefarosa bloqueada con BSA durante 2 horas a 4°C. Eliminamos las bolas de sefarosa y añadimos 3µl del anticuerpo que nos interesa, dejándolo toda la noche en la rueda a 4°C. A la mañana le ponemos de nuevo proteína-A-sefarosa durante 3 horas a 4°C en la rueda, y posteriormente limpiamos estas bolas mediante 5 lavados con tampón RIPA, un lavado con LiCl y dos con TE. Tanto de las muestras como del input guardado eliminamos el RNA mediante incubación con RNAsa durante 30 minutos a 37°C, y extraemos las muestras con NaHCO<sub>3</sub> 0.1M y SDS 1%. Añadimos proteinasa K y dejamos durante toda la noche a 65 °C para hacer el *decrosslinking* y la eliminación de proteínas a la vez.

Finalmente obtenemos el DNA con la técnica de fenol/cloroformo y lo purificamos y limpiamos con etanol, para terminar resuspendiendo en 25µl de H<sub>2</sub>O, de los que utilizaremos 2µl por cada pocillo de la placa de PCR cuantitativa.

### Tampón de lavado A

Hepes pH 7.9 10mM  
EDTA 10mM  
EGTA 0.5mM  
Tritón X100 0.25%

## **Materiales y métodos**

### Tampón de lavado B

Hepes pH 7.9 10mM  
EDTA 1mM  
EGTA 0.5mM  
Tritón X100 0.01%  
NaCl 100mM

### RIPA

NaCl 140mM  
Tris Hcl pH8 10mM  
EDTA 1mM  
SDS 0.1%  
Deoxicolato de sodio 0.1%

### Tampón LiCl

LiCl 250mM  
TrisHCl pH8 10mM  
EDTA 1mM  
NP40 0.5%  
Deoxicolato de Sodio 0.5%

#### 2.4.7.1. Secuenciación y análisis

El DNA inmunoprecipitado con el anticuerpo que reconoce H3T3ph y la muestra input de control se procesa en la *Functional Genomics Core* del Institut de Recerca Biomèdica de Barcelona, y se secuencia en la *Next Generation Sequencing Facility* del *Campus Science Support* (CSF) en Viena, Austria, para la realización de la técnica ChIP-seq. Para ello se generan librerías a partir de 8.5ng de cada una de las muestras una vez cuantificadas mediante *Qubit DNA HS Assay Kit* (Invitrogen) y visualizada la distribución de tamaños de fragmentos del input mediante *Bioanalyzer HS DNA* (Agilent Technologies). La preparación de estas librerías se realiza con el kit *TruSeq RNA Sample Prep kit* (Illumina) y el material generado se amplifica mediante PCR de 14 ciclos siguiendo las instrucciones de la casa comercial. Finalmente se lleva a cabo un control de calidad de las librerías generadas (*Qubit DNA HS* y *Bioanalyzer HS DNA*) para confirmar que tanto la concentración como el tamaño de las librerías generadas son adecuadas para la secuenciación. La modalidad de secuenciación utilizada es 50 nucleótidos simple (50 nt *Single End*) y el proceso de secuenciación se lleva a cabo en un secuenciador *HiSeq 2000* (Illumina).

El preprocesado, control de calidad y análisis bioinformático / bioestadístico de las muestras se realiza en la *Bioinformatics and Biostatistics Core* del Institut de Recerca Biomèdica de Barcelona. En primer lugar se lleva a cabo un análisis general de control de calidad de las muestras mediante el software *FastQC v0.10* (1) para descartar posibles problemas sistemáticos en la secuenciación. Posteriormente se hace el alineamiento de los archivos *FastQ* al genoma de *Drosophila melanogaster* (UCSC dm3) mediante *Bowtie v0.12.5* (Langmead et al., 2009), permitiendo dos *mismatches*, y

reportando solamente alineamientos únicos. Se llevan a cabo controles adicionales de calidad y nivel de enriquecimiento en la inmunoprecipitación de las secuencias ya alineadas mediante *FastQC* y *R v2.15.3/Bioconductor* (2) (Gentleman et al., 2004) utilizando los paquetes *Rsamtools v1.10.2* y *htSeqTools v1.4.0* (3,4). Este último *software* también se usa para detectar y eliminar artefactos de sobreamplificación (mediante la función *filterDuplReads* estableciendo un control de falsos positivos del 1%) y sesgos específicos de cadena (mediante la función *alignPeaks* usando los parámetros por defecto). Se emplea el paquete *rtracklayer v1.18.2* (Lawrence et al., 2009) para generar *tracks* de señal normalizada IP/input a lo largo de todo el genoma.

Para determinar los potenciales sitios de unión se utiliza también el paquete *htSeqTools*, usando la función *enrichedRegions*, seleccionando primero aquellas regiones genómicas cubiertas por un cierto número mínimo de secuencias (utilizando la opción *minReads=10*) y que presentan una acumulación significativa de secuencias en la muestra IP H3T3ph respecto a la muestra input (p-valor ajustado mediante Benjamini-Hochberg  $< 0.05$ ).

Dichas regiones enriquecidas son anotadas al genoma de referencia UCSC dm3 mediante la función *annotatePeakInBatch* del paquete *ChIPpeakAnno* (5). Se utiliza esta información y la distribución general normalizada de señal IP/input a lo largo de todo el genoma para estudiar la colocalización de las regiones enriquecidas de H3T3ph con CP190/Ibf2 (Cuartero et al., 2014), RNA Polimerasa II S2/S5ph (NCBI GEO GSE24115, muestras GSM593411 y GSM593412), H3K27me3 (modENCODE ID 5136), Lamin (NCBI GEO GSE20313, muestras GSM509086 y GSM509086) y CapH2 (NCBI GEO GSE54529, muestras GSM1318355 y GSM1318356), comparando la localización e intensidad de la señal para cada muestra mediante las funciones *regionsCoverage* y *plotMeanCoverage* de *htSeqTools*. Finalmente, se estudia el grado de coincidencia de los genes anotados en las regiones enriquecidas de H3T3ph con los 9 estados de la cromatina observados en la línea celular S2 (Kharchenko et al., 2011), y se lleva a cabo un análisis integrativo general de todas las muestras de interés usando la función *distGPS* del paquete *chroGPS*, utilizando la métrica *iOverlap* (Font-Burgada et al., 2014).

- (1) FastQC 0.10 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- (2) R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>
- (3) Martin Morgan and Hervé Pagès. Rsamtools: Binary alignment (BAM), variant call (BCF), or tabix file import. R package version 1.10.2. <http://bioconductor.org/packages/release/bioc/html/Rsamtools.html>
- (4) Evarist Planet, Camille Stephan-Otto, Oscar Reina, Oscar Flores and David Rossell (2012). htSeqTools: Quality Control, Visualization and Processing for High-Throughput Sequencing data. R package version 1.4.0.
- (5) Lihua Julie Zhu, Herve Pages, Claude Gazin, Nathan Lawson, Jianhong Ou, Simon Lin, David Lapointe and Michael Green (2012). ChIPpeakAnno: Batch annotation of the peaks identified from either ChIP-seq, ChIP-chip experiments or any experiments resulted in large number of chromosome ranges. R package version 2.6.1.



## **Materiales y métodos**

### **2.5. Técnicas de trabajo con lípidos**

#### **2.5.1. Tinción con *Oil Red O***

Disecionamos la glándula protorácica y el cerebro de larvas en estadio III. Los fijamos en paraformaldehído al 4% durante 20 minutos, los lavamos dos veces con PBS durante 10 minutos y los teñimos con una solución de Oil Red O al 0.06% en isopropanol durante 30 minutos. Lavamos las muestras con PBS y las teñimos con DAPI antes de montarlo en Vectashield.

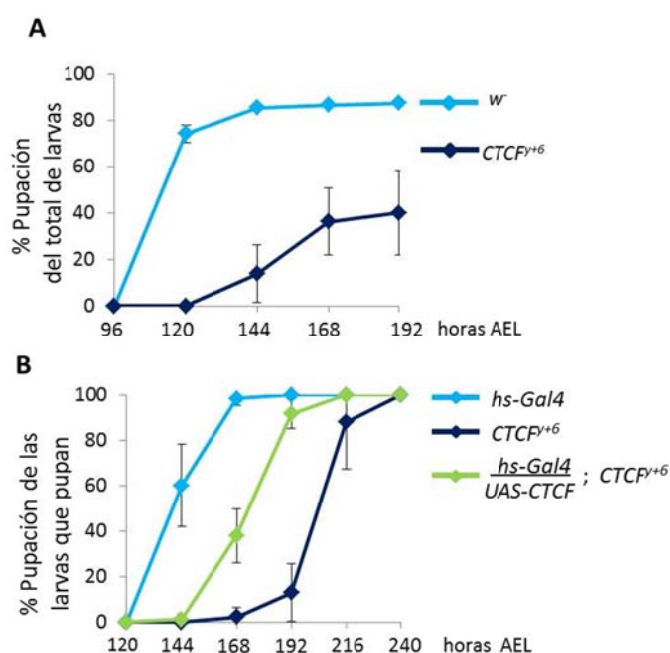
# Resultados



# 1. NUEVAS FUNCIONES DE PROTEINAS *INSULATORS*

## 1.1. CTCF regula el desarrollo en *D. melanogaster*

El análisis del tiempo de desarrollo necesario para darse la transición entre larva y pupa en *Drosophila* muestra que los mutantes nulos para la proteína *insulator* CTCF (*CTCF<sup>y+6</sup>*) (Gerasimova et al., 2007) necesitan de media dos días más que los individuos control (*w*) (Figura 25A). Confirmamos que este retraso se debe efectivamente a la ausencia de esta proteína mediante experimentos de rescate utilizando el sistema *UAS-Gal4*, en el que expresamos CTCF bajo control del promotor del gen *hsp70* (*hs*). De hecho, si a las 24 horas de desarrollo exponemos a los mutantes de esta proteína a 37°C de temperatura durante 20 minutos, el 90% muere antes de pupar, y los que pupan lo hacen con un retraso de 3 días respecto de los control (*hs-Gal4*). Sin embargo, cuando sobreexpresamos CTCF en el fondo mutante, más del 75% consiguen pupar y el retraso en el desarrollo se recupera parcialmente (Figura 25B).



**Figura 25. La ausencia de CTCF produce retraso en el desarrollo.** A/ Porcentaje de larvas control (*w*) y *CTCF<sup>y+6</sup>* que pupan en el tiempo expresado en horas después de la puesta de los huevos (AEL). Las barras de error muestran la desviación estándar de tres experimentos independientes. B/ Porcentaje de larvas control (*hs-Gal4*), *CTCF<sup>y+6</sup>* y *hs>CTCF;CTCF<sup>y+6</sup>* del total de larvas que consiguen pupar respecto al tiempo expresado en horas AEL.

## **Resultados**

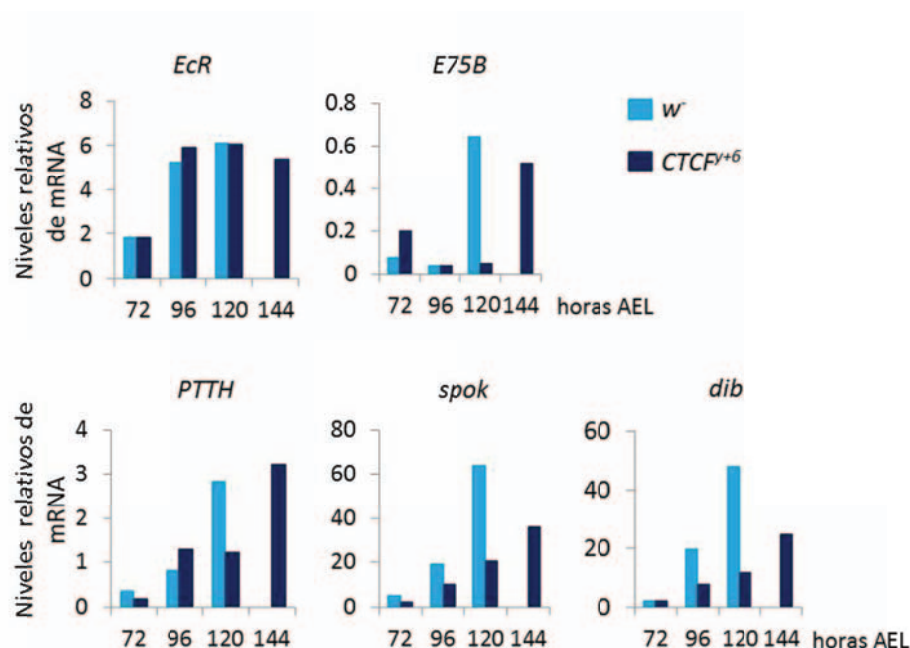
El hecho de realizar la técnica mediante un choque térmico produce algunas variaciones respecto a la primera observación; al exponer los mutantes de *CTCF* a alta temperatura se produce la muerte del 90% de los individuos antes de pupar, mientras que en condiciones normales el 40% consiguen pupar, y el retraso en el desarrollo es en media de más de tres días mientras que es de dos si no se aplica el choque térmico. No obstante, no optamos por otro experimento de rescate que no implique este estrés ya que sobreexpresar la proteína con un *driver* ubicuo produce la muerte a nivel embrionario, y la técnica empleada es suficiente para recuperar el fenotipo como ha sido descrito en otros estudios (Mohan et al., 2007).

### **1.2. CTCF actúa en la vía de síntesis de ecdisona regulando la expresión de los genes *Halloween***

El retraso que acabamos de describir sugiere que la vía de señalización de la hormona ecdisona pueda estar afectada en los mutantes de *CTCF*. Ya que CTCF juega un papel esencial en la regulación de la expresión génica, analizamos los niveles transcripcionales de componentes de esta vía que actúan en diferentes pasos, como por ejemplo el receptor de ecdisona, *E75B*, uno de los genes que responden a esta hormona, *PTTH*, que es la hormona que inicia este proceso, o algunos de los genes *Halloween*, necesarios para la transformación de colesterol en ecdisona, como *spookier (spok)* y *disembodied (dib)*.

Extraemos y purificamos RNA de larvas control (*w*) y *CTCF<sup>y+6</sup>* en diferentes momentos del desarrollo larvario (72, 96, 120 y 144 horas AEL) y, tras retrotranscribirlo a cDNA, comparamos sus niveles mediante PCR cuantitativa.

No observamos ningún cambio en el caso de *EcR*, y para *E75B* y *PTTH* lo único que percibimos es un retraso en la inducción de la expresión de estos genes. Sin embargo, sí que se aprecia una reducción muy notable en los niveles de expresión de los genes involucrados en la biosíntesis de ecdisona, lo que sugiere que CTCF puede jugar un papel en la regulación de la expresión de los genes *Halloween* (Figura 26).

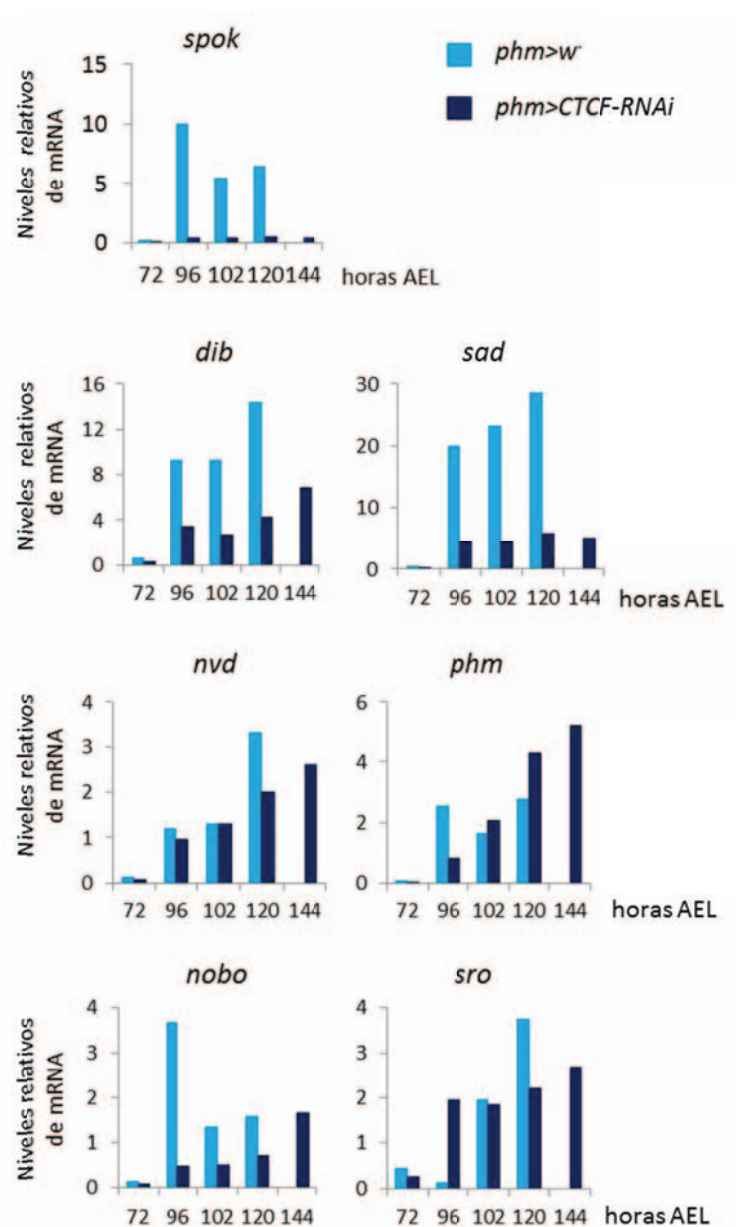


**Figura 26. CTCF regula los niveles de expresión de algunos, pero no todos, de los genes que intervienen en la vía de señalización de la ecdisona.** Los gráficos muestran un experimento representativo de los perfiles transcripcionales normalizado con *RpL23* de los genes indicados en el tiempo expresado en horas AEL tanto en larvas control (*w*) como en *CTCF<sup>y+6</sup>*.

La hormona ecdisona se sintetiza en la glándula protorácica donde los genes *Halloween* son responsables de la transformación de colesterol en ecdisona. Por tanto, analizamos la expresión de estos genes eliminando CTCF únicamente en esta glándula mediante el sistema *UAS-Gal4*, bajo el control del promotor del gen *phantom* (*phm*). Cuantificamos y comparamos los niveles transcripcionales de los genes *Halloween* *spok*, *dib*, *shadow* (*sad*), *neverland* (*nvd*), *phm*, *noppera-bo* (*nobo*) y *shroud* (*sro*) de larvas control (*phm>w*) y de las que carecen de CTCF en la PG (*phm>CTCF-RNAi*).

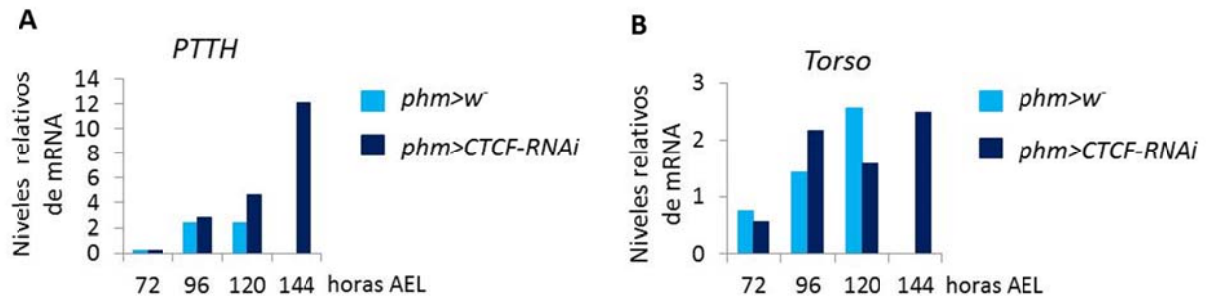
En las larvas control vemos un claro aumento de la inducción de estos genes durante el desarrollo larvario, siendo los niveles bajos al principio e incrementándose a lo largo del tiempo. En las larvas que no tienen CTCF se ven niveles entre 2 y 5 veces menores para *dib*, *sad* y *nobo* en los tiempos de máxima inducción, y no vemos inducción de *spok*. Sin embargo, no hay diferencias claras respecto a otros genes como *nvd*, *phm* y *sro*, lo que nos lleva a pensar que CTCF puede ser necesaria para la activación de varias, pero no de todas, las enzimas involucradas en la síntesis de ecdisona (Figura 27).

## Resultados



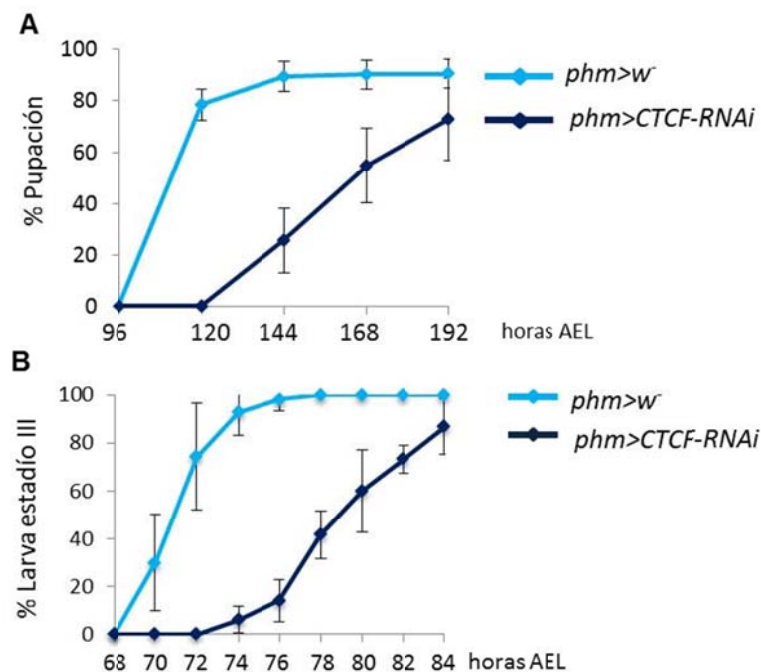
**Figura 27. CTCF regula la expresión de los genes *Halloween* en la glándula protorácica.** Los gráficos muestran un experimento representativo normalizado con *Rpl23* de los perfiles transcripcionales respecto al tiempo expresado en horas AEL de los genes *Halloween* *spok*, *dib*, *sad*, *nvd*, *phm*, *nobo* y *sro* en larvas control (*phm>w*) y *phm>CTCF-RNAi*.

Analizamos si el efecto puede ser indirecto afectando a la hormona que inicia el proceso, *PTTH* (Figura 28A), o al receptor de la misma, *Torso* (Figura 28B), pero no observamos una disminución en los niveles de expresión de estos genes en ausencia de CTCF.



**Figura 28. CTCF no afecta a la expresión de los genes *PTTH* ni *Torso*.** Experimento representativo de los perfiles transcripcionales de *PTTH* (A) y *Torso* (B) normalizado con *Rpl23* a lo largo del tiempo expresado en horas AEL en las larvas control (*phm>w*) y *phm>CTCF-RNAi*.

Observamos también que al disminuir los niveles de CTCF en la PG se produce un retraso en el desarrollo y vemos que la diferencia es de dos días (Figura 29A). También observamos que la transición de larva en estadio II al III se retrasa unas 8 horas (Figura 29B).

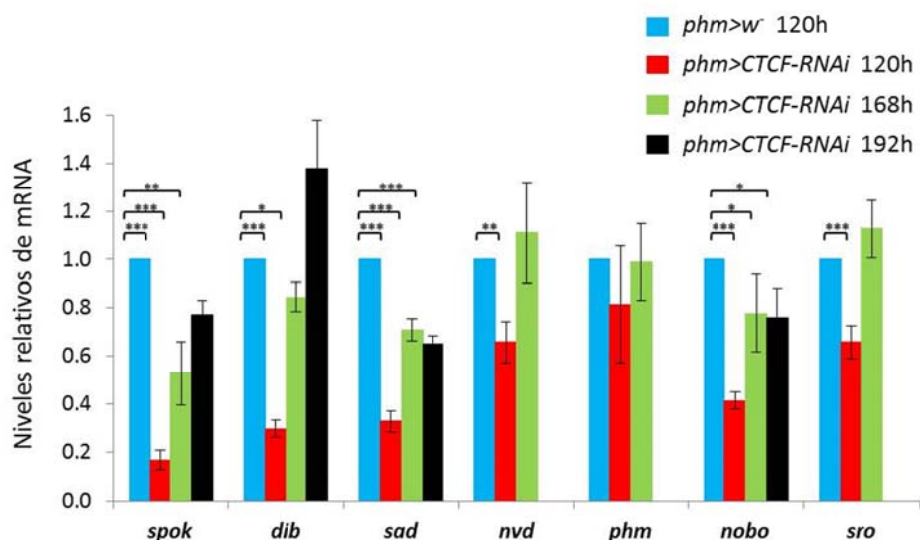


**Figura 29. Retraso en el desarrollo de larvas que no expresan CTCF exclusivamente en glándula protorácica.** A/ Porcentaje de larvas control (*phm>w*) y *phm>CTCF-RNAi* que pupan respecto al tiempo en horas AEL. Las barras de error muestran la desviación estándar de tres experimentos independientes. B/ Porcentaje de larvas de cada genotipo que hacen la transición de larva de estadio II al III respecto al tiempo en horas AEL. Las barras de error muestran la desviación estándar de al menos tres experimentos independientes.



## Resultados

Ya que el desarrollo larvario es 48 horas más largo al no tener CTCF, analizamos la expresión de los genes *Halloween* a las 168 y 192 horas AEL. Vemos que los niveles de *dib*, *nvd*, *phm* y *sro* alcanzados son los mismos que en las larvas control. A pesar de que se observa un incremento en la expresión de *spok*, *sad* y *nobo* al final del desarrollo larvario, los niveles alcanzados son claramente inferiores cuando disminuimos los niveles de CTCF (Figura 30).

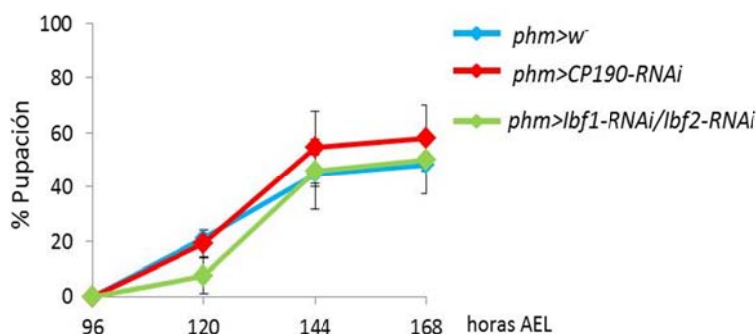


**Figura 30. Niveles transcripcionales de los genes *Halloween*.** Niveles transcripcionales de los genes *Halloween* *spok*, *dib*, *sad*, *nvd*, *phm*, *nobo* y *sro* en larvas *phm>CTCF-RNAi* respecto a las control (*phm>w* 120 h) en los tiempos indicados. Las barras de error muestran los errores estándar de la media (SEM) de al menos tres experimentos independientes. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.1$  según el t test de Student.

### 1.3. Esta función de CTCF es independiente de otras proteínas insulators

Como CTCF forma complejos proteicos con CP190, Ibf1 e Ibf2 analizamos qué ocurre en ausencia de estas proteínas. Para ello, activamos con *phm-Gal4* el RNAi de *CP190* y de *Ibf1* e *Ibf2*, así como el control (*w*), y miramos si hay retraso como previamente hemos señalado para CTCF.

No observamos ningún cambio en el tiempo de desarrollo respecto de los control, lo que sugiere que la regulación del tiempo de desarrollo realizado por CTCF es independiente de los complejos *insulators* (Figura 31).

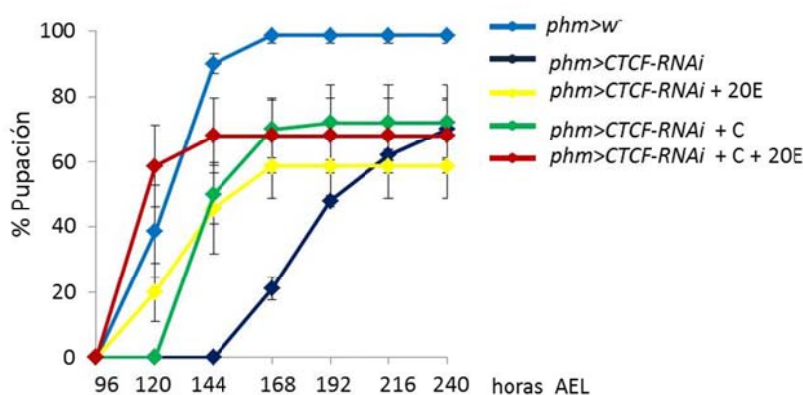


**Figura 31. Tiempo de desarrollo de larvas que no expresan CP190 o Ibf1 e Ibf2 en glándula protorácica.** Porcentaje de larvas control (*phm>w*), *phm>CP190-RNAi* y *phm>lbf1-RNAi/lbf2-RNAi* que pupan respecto al tiempo expresado en horas AEL. Las barras de error muestran la desviación estándar de tres experimentos independientes.

#### 1.4. CTCF es necesaria para la homeostasis del colesterol y síntesis de ecdisona

Para comprobar si lo que produce el retraso en el desarrollo es niveles bajos de ecdisona, alimentamos las larvas con 20-hidroxiecdisona (20E), que es la forma activa de la hormona, o con colesterol, que es el precursor de toda su vía de síntesis, y miramos si hay diferencias en el tiempo de desarrollo.

Aunque algunas de las larvas que no expresan *CTCF* en su PG comienzan a pupar al mismo tiempo que los animales control si les administramos 20E, lo hacen en un período más largo, las pupas son anormales y no se obtienen adultos. Por otro lado, las larvas que han sido alimentadas con colesterol son capaces de pupar correctamente, pero con un retraso de un día respecto a las control. Sin embargo, cuando alimentamos con las dos cosas al mismo tiempo recuperamos completamente el retraso en el desarrollo sugiriendo que *CTCF* es necesaria para la síntesis de ecdisona a partir de colesterol, pero también para la homeostasis del colesterol en las células de la PG (Figura 32).

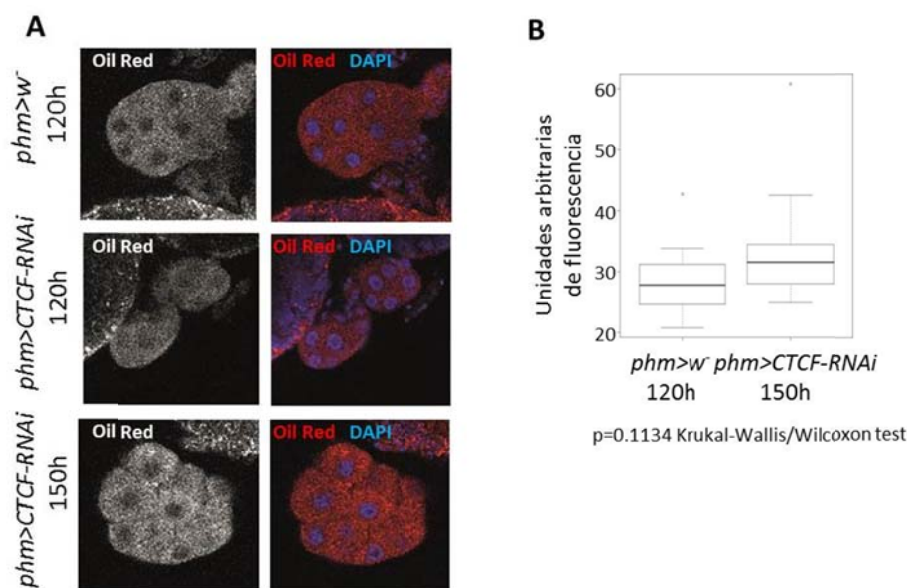


**Figura 32. Ambos, 20-hidroxiecdisona (20E) y colesterol (C) son necesarios para recuperar el fenotipo que produce la ausencia de CTCF en la glándula protorácica.** Tanto las larvas control (*phm>w*) como las *phm>CTCF-RNAi* son alimentadas con comida que contiene etanol, 20E y/o C y se representa el porcentaje de larvas de cada genotipo que pupa respecto al tiempo expresado en horas AEL. Las barras de error muestran las desviaciones estándar de tres experimentos independientes.

## Resultados

Se sabe que en el momento exacto en el que se efectúa la pupación, 120h AEL, se produce una acumulación de gotas de lípidos en las células de la PG, que se pueden detectar mediante la tinción con Oil Red O (Talamillo et al., 2013).

Disecionamos PG de larvas control (*phm>w*) y *phm>CTCF-RNAi* y las teñimos con Oil Red O. En las larvas que no expresan *CTCF* en la PG no vemos esta acumulación a las 120h, probablemente porque en este momento las larvas *phm>CTCF-RNAi* no están lo suficientemente desarrolladas para pupar. Sin embargo, cuando lo estudiamos en el momento en que éstas pupan, sí que lo detectamos, e incluso se observa un pequeño aumento en la acumulación de lípidos respecto a los control (Figura 33).



**Figura 33. Acumulación de lípidos en la glándula protorácica, en larvas control (*phm>w*) y *phm>CTCF-RNAi*.** A/Se muestran gotas de lípidos teñidas con Oil Red O (en blanco y negro, en los paneles de la izquierda, y en rojo, en los de la derecha) a las 120h AEL, tanto de las larvas control (*phm>w*) (arriba) como *phm>CTCF-RNAi* (medio) y para éstas últimas a las 150 horas (abajo). DNA está marcado con DAPI (azul). B/ Gráfico Box plot mostrando la cuantificación de la intensidad de la fluorescencia de Oil Red O de las células de PG de *phm>w* a las 120 horas y *phm>CTCF-RNAi* a las 150 horas.

Todos estos resultados muestran que la expresión de *CTCF* es necesaria en la glándula protorácica para la completa activación transcripcional de los genes *Halloween spookier*, *shadow* y *noppera-bo*, que codifican las enzimas responsables de la biosíntesis de ecdisona, así como para la correcta expresión en el tiempo de los genes que responden a ecdisona. La pérdida de *CTCF* genera un retraso y desincronización en el desarrollo larvario que solo puede ser recuperado alimentando las larvas con la forma activa de la ecdisona y con colesterol. Además, la falta de *CTCF* en las células de la glándula protorácica provoca una acumulación de lípidos. Por todo ello, podemos concluir que *CTCF* es necesaria para la expresión de los genes *Halloween* y la homeostasis del colesterol en las células productoras de ecdisona.

## **2. CARACTERIZACIÓN DE PROTEÍNAS QUE INTERERACTÚAN CON LA PROTEÍNA *INSULATOR* CP190**

### **2.1. Análisis funcional de *lbf1* e *lbf2***

#### **2.1.1. Caracterización de *lbf1* e *lbf2***

En el laboratorio se analizó la composición de los complejos que contienen CP190 y hemos identificado dos nuevas proteínas de función hasta el momento desconocida: *lbf1* e *lbf2*.

Son proteínas codificadas por dos genes del brazo derecho del cromosoma 3, y se encuentran uno al lado del otro. Se transcriben en sentidos opuestos y sus TSS están separados por 470pb.

El gen *lbf1* (3R:5.084.691-5.086.122) tiene 958pb y genera un solo transcrito con cuatro intrones. Da lugar a una proteína de 242 aminoácidos con un peso predicho de 27.7KDa. Se expresa principalmente en embriones de entre 2 y 10 horas. En adultos los máximos niveles de expresión están en el sistema nervioso central y ovarios, a pesar de ser baja comparada con embriones.

El gen *lbf2* (3R:5.083.449-5.084.223) mide 708pb y también da un solo transcrito, pero esta vez con solo un intrón. La proteína que produce tiene 195 aminoácidos, y un peso predicho de 21.9KDa. Al igual que *lbf1* se expresa principalmente en embriones, y en adulto los niveles máximos de expresión están en sistema nervioso central, ovarios y testículos.

Ambas proteínas se parecen y comparten un 18.7% de similitud. Se observa sobre todo una región bien conservada hacia la mitad del dominio N-terminal, que coincide justamente con un *zinc finger* BED, que es un dominio de unión al DNA, y otra en el C-terminal. Además, las dos están conservadas dentro del género *Drosophila*.

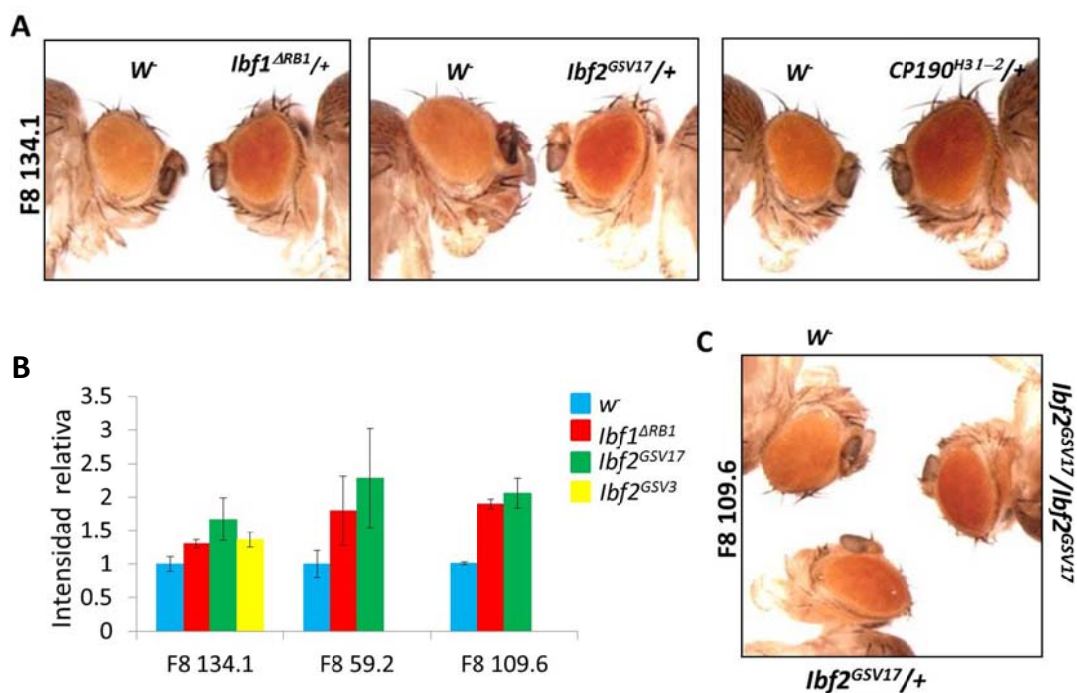
#### **2.1.2. *lbf1* e *lbf2* son proteínas con actividad *enhancer-blocking***

Queremos saber si estas proteínas, que forman un complejo estable con CP190, participan en la función *insulator* y para comprobarlo hacemos ensayos de actividad *enhancer-blocking* con los mutantes de los que disponemos. Ya que hemos observado en experimentos de ChIP-seq que tanto *lbf1* como *lbf2* tienen una fuerte unión al *insulator* Fab8 del complejo Bithorax (Cuartero et al., 2014) caracterizamos su función sobre el mismo. Para realizar el ensayo usamos diferentes líneas transgénicas (F8 134.1, F8 59.2 y F8 109.1) que contienen este *insulator* localizado entre el *enhancer* y el promotor del gen *white* (cuya expresión dota de coloración a los ojos), de manera que dificulta la comunicación entre el *enhancer* y el promotor, y por consiguiente, la expresión del gen. Realizamos un cruce genético de estas líneas con las de los mutantes de *lbf1*, *lbf2* y CP190 (*lbf1*<sup>ΔRB1</sup>, *lbf2*<sup>GSV17</sup>, *lbf2*<sup>GSV3</sup> y CP190<sup>H31-2</sup>) así como con moscas control (*w*).

## Resultados

En los descendientes del cruce de estas líneas con F8 134.1, observamos un efecto muy claro tanto para mutantes de *lbf1* como para los que lo son de *lbf2*, ya que los animales heterocigotos *lbf1*<sup>ΔRB1</sup> e *lbf2*<sup>GSV17</sup> presentan un incremento en la pigmentación de los ojos respecto a los control similar al observado con el mutante de *CP190*, *CP190*<sup>H31-2</sup> (Figura 34A). Además, obtenemos el mismo resultado con otra línea mutante de *lbf2*, *lbf2*<sup>GSV3</sup>, así como con otras líneas transgénicas que contienen Fab8, como F8 59.2 y F8 109.6. Puesto que en estos ensayos se observa cierta variabilidad en la pigmentación de los ojos en diferentes individuos del mismo genotipo cuantificamos la intensidad de la coloración para calcular el valor medio. Corroboramos que efectivamente existe una diferencia en la intensidad del color (Figura 34B). De esta manera confirmamos que estas dos proteínas presentan una actividad característica de las proteínas *insulators*.

Además, un pequeño pero significativo incremento en el color de los ojos se percibe en este ensayo cuando la mutación *lbf2*<sup>GSV17</sup> está en homocigosis respecto a la heterocigota (Figura 34C).



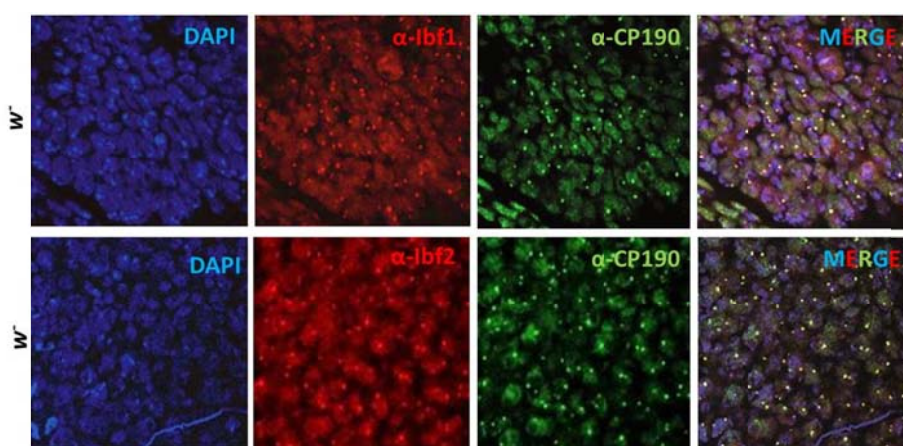
**Figura 34. *lbf1* e *lbf2* tienen actividad *enhancer-blocking* en Fab8.** **A/** Color de ojos de moscas representativas obtenidas del cruce genético de la línea transgénica F8 134.1 con moscas control (*w*<sup>-</sup>) o líneas mutantes (*lbf1*<sup>ΔRB1</sup>, *lbf2*<sup>GSV17</sup> o *CP190*<sup>H31-2</sup>), presentando la mutación en heterocigosis. **B/** Intensidad relativa de la pigmentación de los ojos de moscas obtenidas del cruce genético de las líneas transgénicas F8 indicadas con líneas mutantes (*lbf1*<sup>ΔRB1</sup>, *lbf2*<sup>GSV17</sup> o *lbf2*<sup>GSV3</sup>) respecto a las obtenidas del cruce con las control (*w*<sup>-</sup>), en heterocigosis. Las barras de error muestran las SEM de 3 experimentos independientes. **C/** Color de ojos de moscas representativas obtenidas del cruce genético de la línea transgénica F8 109.6 con moscas control (*w*<sup>-</sup>) o el mutante *lbf2*<sup>GSV17</sup>, presentando la mutación en heterocigosis y homocigosis.

Con estos resultados podemos concluir que *lbf1* e *lbf2* tienen actividad *enhancer-blocking* en el *insulator* Fab8 del complejo Bithorax.

### 2.1.3. Ibf1 e Ibf2 forman *insulator bodies*, y se necesitan mutuamente

Las diferentes proteínas *insulators* pueden interaccionar entre ellas a pesar de que se encuentren unidas al DNA en secuencias muy alejadas en el genoma. Su colocalización puede verse como unos corpúsculos nucleares denominados *insulator bodies*. Para comprobar si Ibf1 e Ibf2 están presentes en los *insulator bodies* hacemos inmunotinciones de células de disco imaginal de ala de larvas control con los anticuerpos específicos contra estas dos proteínas. Además, ya que han sido identificadas como proteínas que forman un complejo estable con CP190, miramos si cada una de ellas colocaliza con CP190.

En la figura se observa que tanto Ibf1 como Ibf2 localizan en unos corpúsculos dentro del núcleo, que además, la mayoría de ellos, coinciden con los de CP190, indicando que las tres proteínas colocalizan en *insulator bodies* (Figura 35).

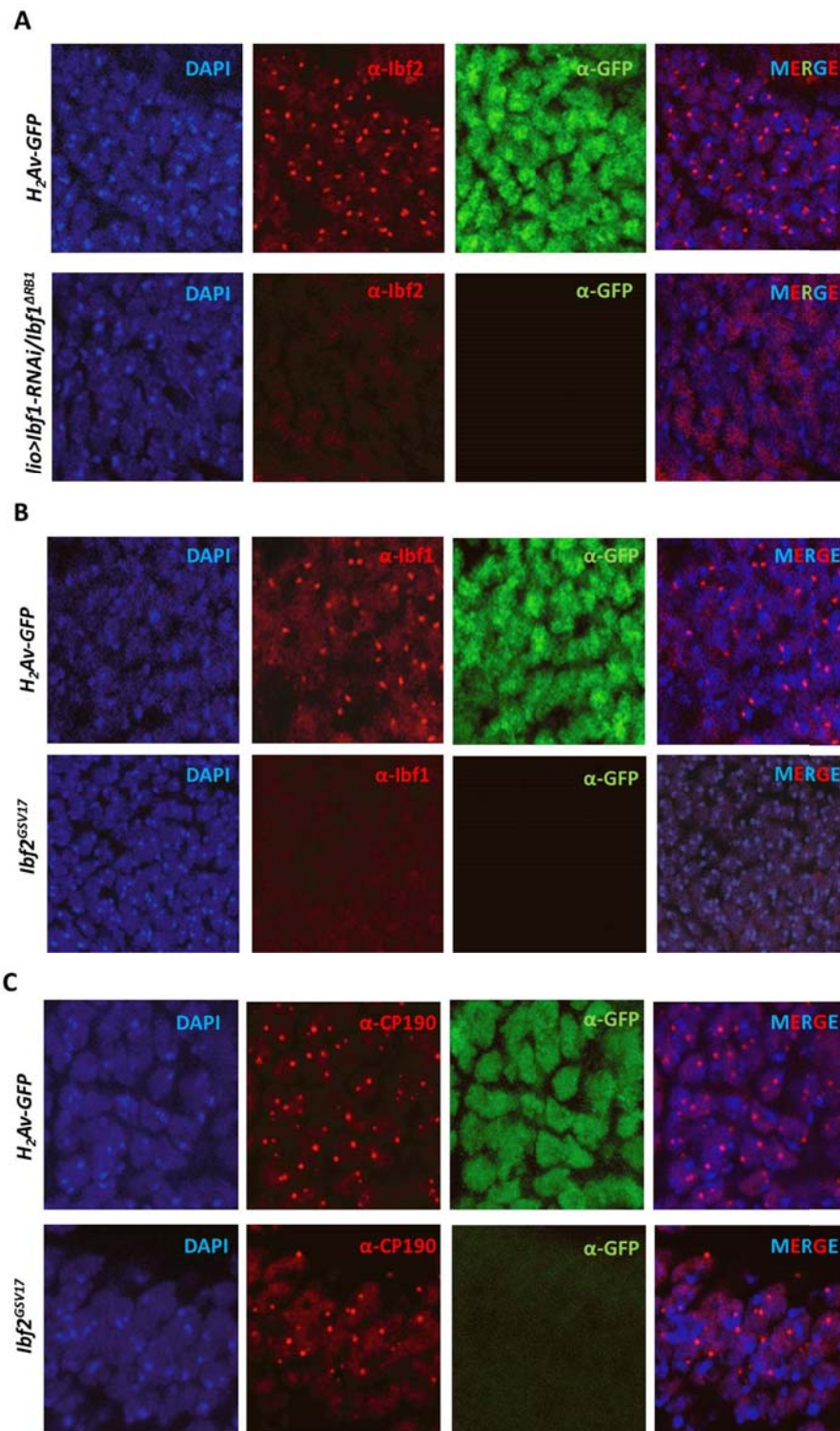


**Figura 35. Ibf1 e Ibf2 forman *insulator bodies* y colocalizan con CP190.** Immunolocalización de Ibf1 (rojo, panel superior) o Ibf2 (rojo, panel inferior) y CP190 (verde) en células de disco imaginal de ala en individuos control (w). DNA está teñido con DAPI (azul).

Nos preguntamos si cada una de ellas necesita a alguna de las otras para localizar en estos cuerpos, así que miramos su presencia en larvas que no expresan alguna de las otras dos proteínas. Utilizamos como larvas control una línea que tiene unido el tag GFP a la H2Av (Clarkson and Saint, 1999), de manera que mediante la detección del GFP podemos distinguir aquellas que son control de las mutantes. Mientras que las larvas mutantes nulas para *Ibf2* llegan a estadio III, para conseguir la ausencia de Ibf1 utilizamos larvas que combinan la mutación de *Ibf1* con un RNAi de *Ibf1* que se expresa mediante el sistema *UAS-Gal4* bajo control del promotor del gen *lto*, ya que las larvas homocigotas mutantes de *Ibf1* no llegan a larva III. Las larvas mutantes de *CP190* son larvas que combinan dos mutaciones diferentes de *CP190*, *CP190<sup>P1</sup>/CP190<sup>P11</sup>*, ya que cada una de ellas no es viable en homocigosis hasta este punto del desarrollo (Pai et al., 2004).



## Resultados



**Figura 36. Ibf1 e Ibf2 se necesitan para localizar en *insulator bodies*, pero CP190 no depende de ellas. A/** Immunolocalización de Ibf2 (rojo) en larvas control (*H2Av-GFP*) (panel de arriba, GFP detectado en verde) y mutantes de *Ibf1* (*lio>Ibf1-RNAi/Ibf1<sup>ΔRB1</sup>*) (panel de abajo). DNA está teñido con DAPI (azul). **B/** Immunolocalización de Ibf1 (rojo) en larvas control (*H2Av-GFP*) (panel de arriba, GFP detectado en verde) y mutantes de *Ibf2* (*Ibf2<sup>GSV17</sup>*) (panel de abajo). DNA está teñido con DAPI (azul). **C/** Immunolocalización de CP190 (rojo) en larvas control (*H2Av-GFP*) (panel de arriba, GFP detectado en verde) y mutantes de *Ibf2* (*Ibf2<sup>GSV17</sup>*) (panel de abajo). DNA está teñido con DAPI (azul).

Observamos que en ausencia de Ibf1, Ibf2 no forma *insulator bodies* (Figura 36A). Cuando lo hacemos a la inversa, vemos que la carencia de la segunda impide la detección de la primera en estos cuerpos (Figura 36B). Sin embargo, la ausencia de Ibf2 (y por tanto de Ibf1) no afecta a que CP190 se detecte en los *insulator bodies* (Figura 36C).

Con estos resultados podemos concluir que Ibf1 e Ibf2 localizan en *insulator bodies* junto con CP190, y que ambas proteínas se necesitan mutuamente para ello. En cambio, la localización de CP190 no depende de las otras proteínas.

## **2.2. Caracterización y análisis funcional de pita**

### **2.2.1. Características de pita**

Otra de las proteínas identificadas en el complejo de CP190 es pita. Según la base de datos Flybase, el gen *pita*, CG3941, localizado en el cromosoma 2R (2R:23548967..23554712), anida al de la caspasa Dcp-1. Da lugar a dos transcritos y dos polipéptidos, de 683 y 623 aminoácidos. Se expresa principalmente en embriones de entre 0 y 10 horas, así como en sistema nervioso central y glándulas salivares en fase larvaria, y en ovarios y glándulas salivares en adulto.

*Pita* da lugar a una proteína nuclear con capacidad de unión al DNA gracias a sus *Zinc fingers*. Regula la replicación del DNA, y su falta produce hipercondensación de cromosomas y rotura de los mismos en células mitóticas. Su ausencia provoca una acumulación de células con un contenido en DNA entre 2N y 4N, característico de un aumento de células en fase S, en tejidos proliferativos y también afecta a la endorreduplicación en glándulas salivares. Se ha descrito como factor de transcripción que activa la expresión de genes, como *Orc4*, requerido para la iniciación de la replicación (Page et al., 2005).

La ausencia de pita es letal y en línea germinal produce defectos en el desarrollo de las cámaras ováricas, con núcleos de las células nodrizas anormales o ausentes, que conlleva la muerte celular prematura (Laundrie et al., 2003).

### **2.2.2. Pita forma parte del complejo CP190-Ibf1-Ibf2**

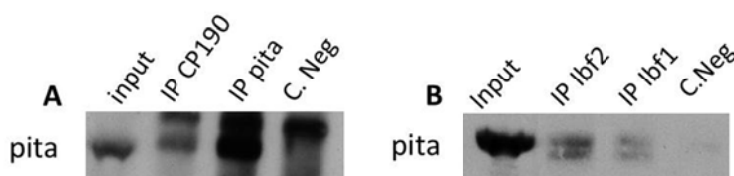
En primer lugar verificamos la interacción entre CP190 y pita mediante ensayos de coimmunoprecipitación en células S2. Immunoprecipitamos la muestra con los anticuerpos contra CP190 y pita, y mediante *Western Blot* detectamos pita en ambos casos (Figura 37A). Por tanto podemos concluir que CP190 y pita están en el mismo complejo.

Una vez comprobada la interacción con CP190, nos preguntamos si forma parte del mismo complejo que engloba a CP190, Ibf1 e Ibf2, así que realizamos immunoprecipitaciones con los



## Resultados

anticuerpos específicos para estas dos proteínas, y hacemos el *Western blot* para detectar la presencia de *pita* dentro del grupo de proteínas que inmunoprecipitamos con ellas (Figura 37B). Nuestros resultados muestran que *pita* coimmunoprecipita con *Ibf1* e *Ibf2*, indicando que CP190, *Ibf1*, *Ibf2* y *pita* forman un complejo estable en *Drosophila*.

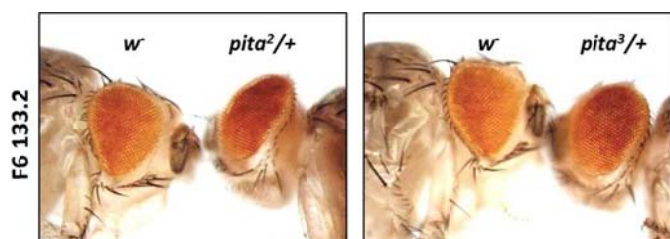


**Figura 37. Pita coimmunoprecipita con CP190, *Ibf1* e *Ibf2*.** A/ *Western blot* revelado contra *pita* de extractos nucleares de células S2 que se han inmunoprecipitado con los anticuerpos indicados ( $\alpha$ -CP190,  $\alpha$ -*pita*) o sin anticuerpo (C. Neg). Input corresponde al 10% del material inmunoprecipitado. B/ *Western blot* revelado contra *pita* de extractos nucleares de células S2 que se han inmunoprecipitado con los anticuerpos indicados ( $\alpha$ -*Ibf2*,  $\alpha$ -*Ibf1*) o sin anticuerpo (C. Neg). Input corresponde al 10% del material inmunoprecipitado.

### 2.2.3. Pita tiene actividad *enhancer-blocking*

Analizamos la actividad *insulator* de *pita* en ensayos de *enhancer-blocking*. En este caso utilizamos una línea transgénica con el *insulator* Fab6, F6 133.2, entre *enhancer* y promotor del gen *white*. Usamos dos mutaciones de *pita* diferentes, *pita*<sup>2</sup> y *pita*<sup>3</sup> (Laundrie et al., 2003). *pita*<sup>2</sup> es letal a partir de estadio prepupa mientras que *pita*<sup>3</sup> lo es ya desde larva.

Como se aprecia en las imágenes, los descendientes obtenidos de los cruces genéticos realizados con la línea F6 133.2 y ambas líneas mutantes de la proteína presentan un incremento en la pigmentación de la coloración de los ojos respecto a los descendientes del cruce realizado con individuos control, lo que sugiere que *pita* pueda tener actividad *enhancer-blocking* en Fab6 (Figura 38). Lo hemos analizado también para una línea transgénica que contiene Fab7, F7 B+P2.5, pero no hemos visto diferencias.



**Figura 38. Pita es necesaria para la actividad *enhancer-blocking* en Fab6.** Color de ojos de moscas representativas obtenidas del cruce genético de la línea transgénica F6 133.2 con moscas control (*w*) o mutantes de *pita* (*pita*<sup>2</sup> (izquierda) o *pita*<sup>3</sup> (derecha)), presentando la mutación en heterozigosis.

Nuestros resultados sugieren que *pita* es necesaria para la actividad del elemento Fab6 pero es necesario completar estos estudios utilizando otras líneas transgénicas para saber si realmente la proteína desempeña esta actividad *insulator*.

### **3. NUEVAS PROTEINAS CON ACTIVIDAD *INSULATOR*: Haspin**

#### **3.1. Identificación de Haspin como proteína *insulator***

Para buscar nuevas proteínas con actividad *insulator* realizamos una mutagénesis en *Drosophila* mediante inserción de un elemento P y analizamos cada una de las líneas obtenidas buscando cambios en la expresión del gen *reporter white* en ensayos de *enhancer-blocking*, utilizando la construcción que contiene el *insulator* del Fab7 entre el *enhancer* y el promotor del gen *white*.

El elemento P es un transposón muy utilizado en *Drosophila* como mutágeno. Para evitar la transposición incontrolada del elemento se han creado líneas transgénicas que contienen, o las secuencias que codifican para la enzima transposasa, o las secuencias necesarias para que esta enzima reconozca el elemento y pueda escindirlo y reinsertarlo de manera aleatoria en cualquier otro lugar del genoma. De esta manera el salto y re inserción del transgén solo se da al cruzar las dos líneas, y las nuevas inserciones quedan fijadas en la descendencia al eliminar la transposasa. Cada una de ellas puede entonces ser analizada fenotípicamente, en nuestro caso observando el color de los ojos y mantenida gracias a la presencia en el transgén del gen *reporter yellow*. Analizamos unos 450 saltos diferentes y en 27 observamos cambios en la expresión del gen *reporter white*. Mediante ensayos de PCR-inversa caracterizamos la localización en el genoma de estas nuevas inserciones y continuamos nuestros estudios con una de ellas, la línea que tenía insertado el transgén en la región 5' no traducida (UTR) del gen *Haspin*.

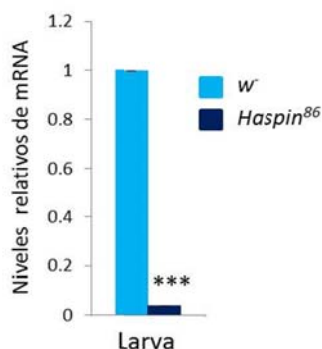
#### **3.2. Haspin**

Haspin es una proteína conservada en eucariotas (Higgins, 2001), pero no ha sido hasta el momento estudiada en *Drosophila*. La única información de la que se dispone es de la que está recogida en la base de datos Flybase. Según la información que aquí hay sobre el gen *Haspin*, nombrado con el símbolo Dmel\Haspin (CG40080, FBgn0046706), fue descrito por primera vez en *Drosophila* en 2001 por Manning y codifica para una proteína de localización nuclear, con actividad kinasa sobre serinas y treoninas. Tiene regiones no transcritas tanto en el extremo 5' como en el 3', y presenta 5 exones y 4 intrones. Solo tiene un transcrito de 1815 nucleótidos, y produce un único polipéptido de 566 aminoácidos. Se localiza en la heterocromatina del cromosoma 2R, concretamente 2R:4014111..4049342. La proteína forma una subfamilia dentro de las Ser/Thr kinasas, ya que presenta un dominio kinasa particular.

El perfil de expresión temporal del gen, de acuerdo a lo publicado en FlyBase, varía desde niveles bajos hasta otros moderadamente altos, observándose el máximo en embriones de entre 0 y 18 horas.

## Resultados

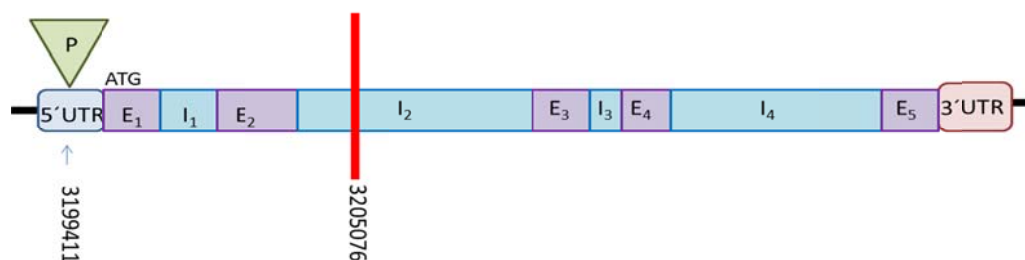
No existen mutantes de *Haspin* ni inserciones de elementos P en la región del gen *Haspin*, así que caracterizamos la línea obtenida en nuestro *screening*, *Haspin*<sup>86</sup>. Medimos el nivel de transcripción del gen de individuos control (*w*) y *Haspin*<sup>86</sup> a través de la extracción de mRNA de larvas enteras y su posterior retrotranscripción a cDNA, mediante PCR cuantitativa. Comparando los resultados alcanzados en los individuos control y en la línea *Haspin*<sup>86</sup> en homocigosis, detectamos una bajada de más del 90% en la transcripción del gen (Figura 39). Por tanto, podemos concluir que tenemos un mutante nulo o hipomorfo muy fuerte.



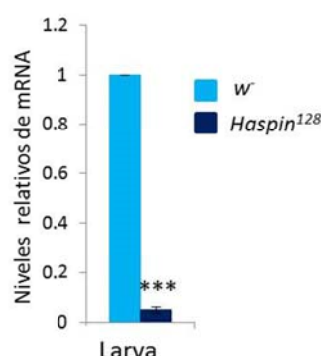
**Figura 39. Disminución del nivel de expresión del gen *Haspin* en la línea *Haspin*<sup>86</sup>.** Representación relativa del nivel transcripcional del gen *Haspin* respecto a *tubulina* en larvas control (*w*) y mutantes de *Haspin* (*Haspin*<sup>86</sup>). Las barras de error muestran los SEM de dos experimentos independientes. \*\*\*  $p < 0.001$  según test t de Student.

A pesar de que esta bajada del nivel transcripcional de *Haspin* tan fuerte en la línea *Haspin*<sup>86</sup> sugiere que la mutación es nula, la única forma de estar seguros de que una línea sea nula para una proteína es mediante la escisión del gen, o una parte del mismo que contenga ATG que codifica para el inicio de la síntesis de la proteína. Para ello provocamos el salto del elemento P cruzando la línea *Haspin*<sup>86</sup> con otra que expresa una transposasa, capaz de hacer saltar el transposón. Este salto puede ser limpio, desplazando solo el elemento P, o bien no serlo, llevándose parte de la secuencia génica con él.

De 60 descendientes en los que se ha producido la escisión del elemento P, analizamos qué tipo de salto ha producido a través de PCR mediante la extracción de DNA de larva entera. De ellos, uno indica que se ha llevado consigo el ATG. Caracterizamos esta línea mediante PCR y secuenciación. Confirmamos que el elemento P se ha llevado 5700pb, dejando a la vez restos de sí mismo (Figura 40). Concretamente, ha eliminado los dos primeros exones, todo el primer intrón y parte del segundo, de manera que se ha producido la escisión del ATG del gen. De esta forma conseguimos un mutante nulo ya que no es capaz de sintetizar la proteína. A esta línea la llamamos *Haspin*<sup>128</sup>. Analizamos con esta nueva línea el nivel transcripcional del gen de igual manera que hemos hecho previamente con *Haspin*<sup>86</sup> y detectamos que disminuye en más del 90% respecto a los control (*w*) (Figura 41).



**Figura 40. Generación de un mutante nulo para *Haspin*.** Representación gráfica del gen *Haspin*, donde “P” indica la posición exacta donde está insertado el elemento P en la línea *Haspin*<sup>86</sup>. Desde su localización hasta la barra roja representa la parte del genoma que se ha llevado en el salto. ATG: metionina que comienza la síntesis de la proteína. E: exón. I: intrón. UTR: región no transcrita del gen en el extremo 5-terminal o 3-terminal.



**Figura 41. Disminución del nivel de expresión del gen *Haspin* en la línea *Haspin*<sup>128</sup>.** Representación relativa del nivel transcripcional del gen *Haspin* respecto a *actina* en larvas control (*w*) y mutantes de *Haspin* (*Haspin*<sup>128</sup>). Las barras de error muestran los SEM de tres experimentos independientes. \*\*\* p<0.001 según test t de Student.

### 3.3. Fenotipos

La alteración de la correcta expresión de los genes homeóticos, como es *Abd-B*, produce transformaciones homeóticas. Nuestro *screening* ha mostrado que *Haspin* es necesaria para la actividad del *insulator* del Fab7 de la región *Abd-B*. Sin embargo los mutantes de *Haspin* no presentan ningún fenotipo de transformación homeótica.

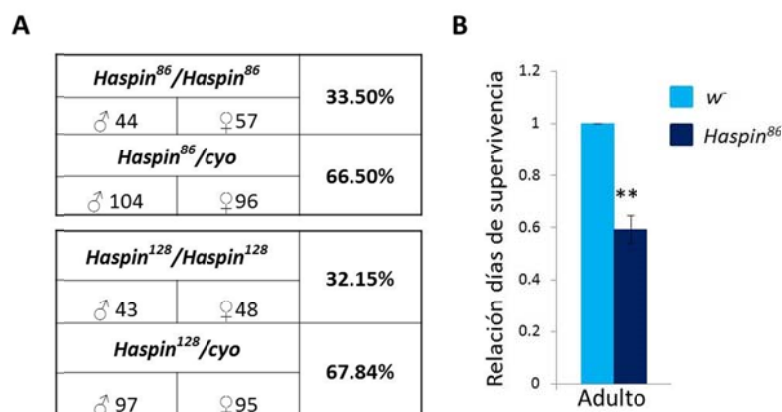
A pesar de que la transcripción del gen *Haspin* está reducida en más del 90%, individuos con la mutación en homocigosis sobreviven hasta fase adulta, tanto en la línea *Haspin*<sup>86</sup> como en *Haspin*<sup>128</sup>.

Analizamos la viabilidad de la descendencia partiendo de padres heterocigotos; la mutación está balanceada con *cyo*, de manera que no puede existir descendencia homocigota para el balanceador. En caso de ser 100% viable, obtendríamos el 33% de descendientes homocigotos y 66% heterocigotos. Para la línea *Haspin*<sup>86</sup> obtenemos 200 descendientes, de los cuales 33.5% son homocigotos y 66.5% heterocigotos (Figura 42A, panel superior). Hacemos de nuevo el estudio con la

## Resultados

línea *Haspin*<sup>128</sup> y de los 283 descendientes analizados, 32.1% son homocigotos frente al 67.54% que presentan la mutación en heterocigosis (Figura 42A, panel inferior).

Hacemos un ensayo de supervivencia con la línea *Haspin*<sup>86</sup> y vemos que sí que está afectada, pues los mutantes homocigotos viven menos días que los individuos control (Figura 42B). La diferencia observada es estadísticamente significativa de acuerdo al test t de Student ( $p=0.00143$ ).



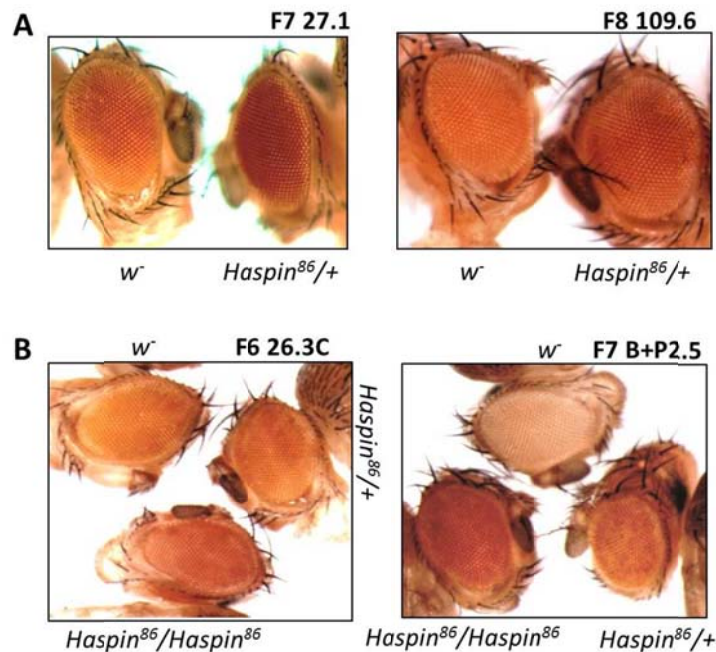
**Figura 42. La ausencia de Haspin no causa letalidad, pero sí afecta a la supervivencia.** A/ Número de descendientes de cada genotipo señalado, así como el porcentaje que representa respecto al total de descendientes analizados en la línea *Haspin*<sup>86</sup> (panel superior) y en la línea *Haspin*<sup>128</sup> (panel inferior). B/ Representación relativa de la duración de la vida de los adultos homocigotos *Haspin*<sup>86</sup> respecto a los control (*w*<sup>-</sup>). \*\*  $p<0.005$  según test t de Student.

Por tanto, podemos concluir que la ausencia de Haspin no afecta a la viabilidad en condiciones de laboratorio. Sin embargo, sí que se observa una disminución del tiempo de vida de los adultos.

### 3.4. Haspin es una proteína con actividad enhancer-blocking

Para caracterizar en profundidad la actividad *insulator* de Haspin, llevamos a cabo ensayos *enhancer-blocking*. Realizamos un cruce genético entre la línea *Haspin*<sup>86</sup> con diferentes líneas transgénicas que tienen insertado un *insulator* entre el *enhancer* y el promotor del gen *white*. Utilizamos líneas transgénicas que contienen los *insulators* Fab6 (F6 34.2 y F6 26.3C), Fab7 (F7 B+P2.5 y F7 27.1) o Fab8 (F8 109.6, F8 134.1 y F8 59.2) entre los elementos reguladores, y en los descendientes de todos estos cruces obtenemos un incremento en la pigmentación de los ojos respecto al mismo cruce realizado con individuos control (*w*<sup>-</sup>) (Figura 43A muestra el ensayo con las líneas F7 27.1 y F8 109.6). Estos análisis se han analizado con la mutación en heterocigosis, y

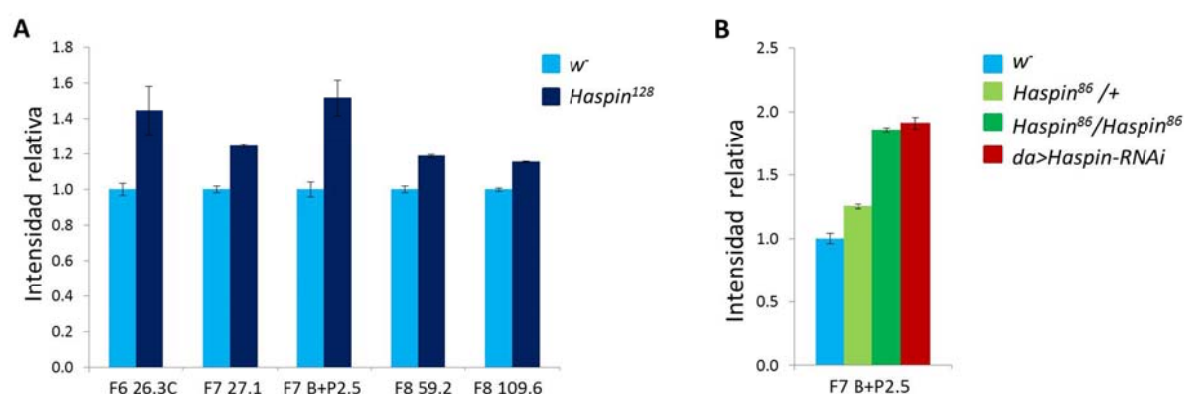
observamos que el efecto se incrementa incluso si se encuentra en homocigosis (Figura 43B muestra el ensayo con las líneas F6 26.3C y F7 B+P2.5).



**Figura 43. Haspin es necesaria para la actividad *enhancer-blocking* en diferentes Fabs del complejo Bithorax. A/** Color de ojos de moscas representativas obtenidas del cruce genético de las líneas transgénicas indicadas en cada imagen con moscas control ( $w^-$ ) o mutantes de *Haspin* ( $Haspin^{86}$ ), presentando la mutación en heterocigosis. **B/** Color de ojos de moscas representativas obtenidas del cruce genético de las líneas transgénicas indicadas en cada imagen con moscas control ( $w^-$ ) o mutantes de *Haspin* ( $Haspin^{86}$ ), presentando la mutación en heterocigosis y homocigosis.

Para descartar que el efecto observado pueda ser debido al fondo genético de la línea *Haspin*<sup>86</sup> realizamos los ensayos con el otro mutante, *Haspin*<sup>128</sup>, así como disminuyendo los niveles de Haspin mediante RNAi y el sistema *UAS-Gal4*, gracias al *driver da-Gal4*. Cuantificamos la intensidad de la pigmentación y comprobamos que la mutación *Haspin*<sup>128</sup> en heterocigosis tiene efecto en la actividad *enhancer-blocking* de los diferentes *insulators* (Figura 44A) y que la disminución de los niveles de Haspin mediante RNAi también tiene un efecto claro en la actividad *insulator* en el Fab7 (Figura 44B).

## Resultados



**Figura 44. Haspin tiene actividad *enhancer-blocking*.** **A/** Intensidad relativa de la pigmentación de los ojos de moscas obtenidas del cruce genético de las líneas transgénicas indicadas con la línea mutante *Haspin*<sup>128</sup> respecto a las obtenidas del cruce con las control (*w*), en heterozigosis. Las barras de error muestran las SEM de 3 experimentos independientes. **B/** Intensidad relativa de la pigmentación de los ojos de moscas obtenidas del cruce genético de la línea transgénica F7 B+P2.5 con las líneas mutantes de *Haspin* indicadas respecto a las obtenidas del cruce con las control (*w*). Las barras de error muestran las SEM de 3 experimentos independientes.

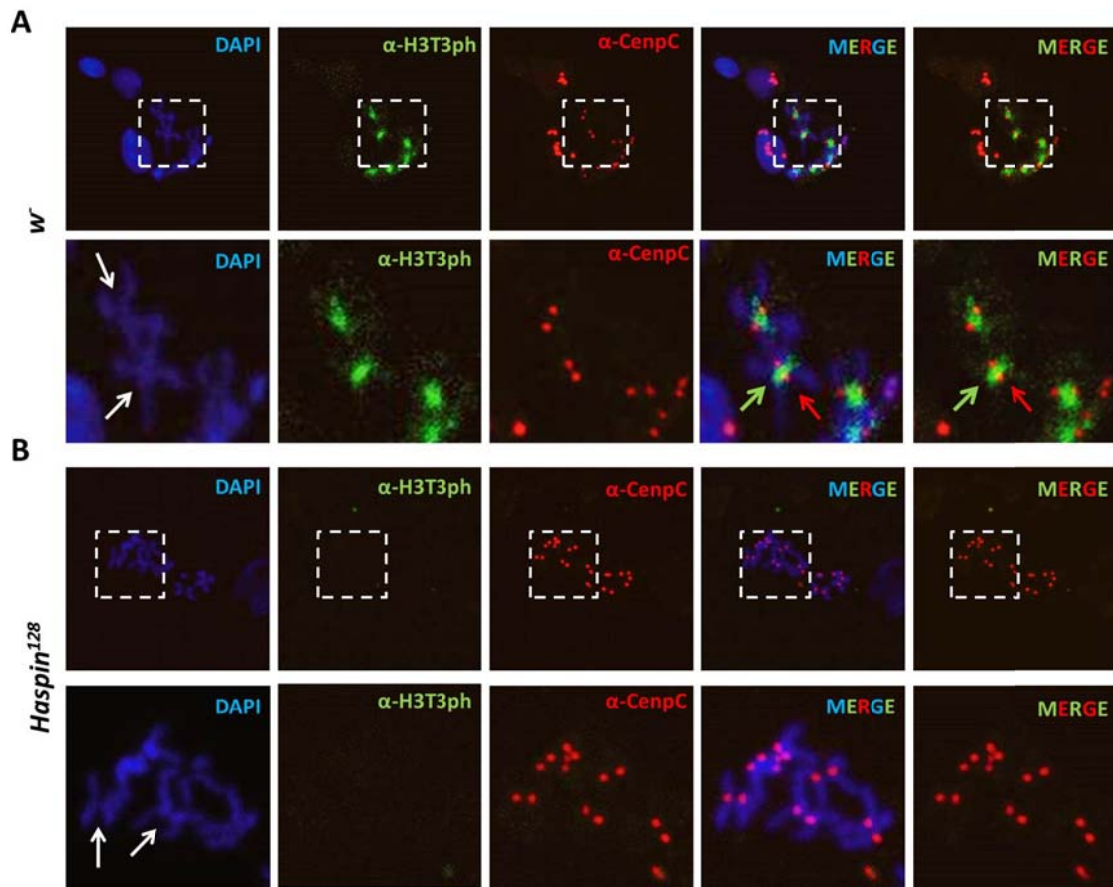
Concluimos que Haspin es necesaria para la actividad *enhancer-blocking* de los *insulators* Fab6, Fab7 y Fab8 de la región reguladora de gen *Abdominal-B*, confirmando que tiene una actividad característica de las proteínas *insulators*.

### 3.5. Actividad kinasa

#### 3.5.1. Mitosis

Haspin ha sido descrita en otros organismos como la kinasa que fosforila la treonina 3 en la histona 3 durante mitosis (Yamagishi et al., 2010, Carretero et al., 2013). Estudiamos si en *Drosophila* también es así mediante un *squash* de células de cerebro de una línea control (*w*). Observamos la presencia de H3T3ph en los cromosomas mitóticos. Vemos dos señales que quedan entre las cromátidas hermanas de cromosomas metafásicos. Mediante la coimmunolocalización con CenpC, que marca centrómeros a lo largo de todo el ciclo, podemos concretar que la vemos en la región pericentromérica, lugar que también ocupa en otros organismos (Figura 45A) (Dai et al., 2005). Hacemos el mismo análisis con la línea mutante de *Haspin* *Haspin*<sup>128</sup> y observamos que mientras que la señal de Cenp-C se mantiene invariable en los centrómeros, la de H3T3ph desaparece. Además, parece apreciarse que la falta de H3T3ph provoca una separación de las cromátidas (Figura 45B).





**Figura 45. Haspin fosforila H3T3 en la región pericentromérica. A/** Immunolocalización de CenpC (rojo) y H3T3ph (verde) en cromosomas mitóticos de células de cerebro en larvas control (*w*). DNA está teñido con DAPI (azul). En el panel inferior se ve una ampliación de la zona recuadrada. Flecha roja señala centrómeros, flecha verde localización de H3T3ph entre las cromátidas hermanas y flecha blanca unión de las cromátidas. **B/** Immunolocalización de CenpC (rojo) y H3T3ph (verde) en cromosomas mitóticos de células de cerebro en larvas mutantes *Haspin*<sup>128</sup>. DNA está teñido con DAPI (azul). En el panel inferior se ve una ampliación de la zona recuadrada. Flecha blanca señala separación de crómatidas.

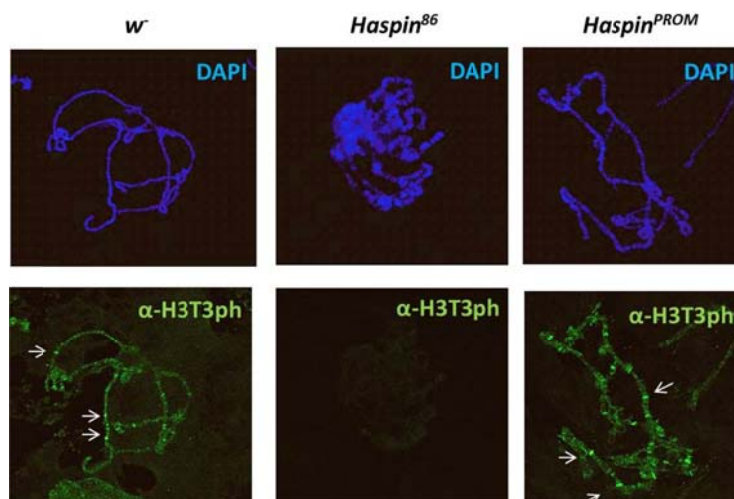
### 3.5.2. Interfase

Se sabe que en mitosis se produce un aumento de la actividad de Haspin y, por tanto, de la fosforilación de H3T3 (Ghenoiu et al., 2013). No obstante, se ha mostrado que el dominio kinasa de Haspin es activo a lo largo de todo el ciclo (Dai et al., 2005, Villa et al., 2009, Higgins, 2010). Para estudiar H3T3ph en interfase hacemos inmunotinción de cromosomas politénicos, específicos de glándulas salivares y otros tejidos de *Drosophila*. Son cromosomas gigantes resultado de múltiples replicaciones de su material genético que no experimentan mitosis. Las cromátidas se mantienen juntas, presentando un patrón de bandas característico, con zonas más oscuras (bandas) donde se sitúa la heterocromatina, y zonas más claras (interbandas) con la cromatina activa. Aunque la señal no es muy potente, observamos presencia de H3T3ph en interbandas (Figura 46, izquierda), y desaparece en su totalidad en la línea *Haspin*<sup>86</sup> (Figura 46, centro).



## Resultados

Creamos una línea que sobreexpresa Haspin bajo el control de su promotor endógeno, *Haspin<sup>PROM</sup>*. Realizamos el mismo experimento sobre ella y se observa un incremento de la señal (Figura 46, derecha).



**Figura 46. H3T3ph está presente en interfase.** Immunolocalización de H3T3ph (verde) en cromosomas politénicos en larvas control (*w<sup>-</sup>*) (izquierda), *Haspin<sup>86</sup>* (centro) y *Haspin<sup>PROM</sup>* (derecha). DNA está teñido con DAPI (azul). Flechas señalan algunas de las bandas de H3T3ph.

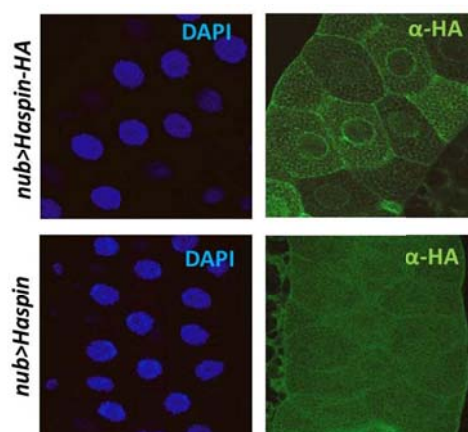
Por tanto, podemos concluir que Haspin es la kinasa que fosforila H3T3 en *D. melanogaster*, no solo durante mitosis, sino que también en interfase.

### 3.6. Organización nuclear

#### **3.6.1. Localización de Haspin en interfase**

Con el fin de caracterizar la proteína en nuestro animal modelo creamos un anticuerpo que nos permita detectarla. Lo hacemos contra la región N-terminal de la proteína, ya que la C-terminal contiene el dominio kinasa, evitando de esta manera reconocimientos inespecíficos de otras proteínas con un dominio similar. Sin embargo, utilizando este anticuerpo no somos capaces de detectar la presencia de la proteína en individuos control ni por *Western blot* ni por inmunotinción.

Por ello, obtenemos una línea transgénica en la que se expresa Haspin fusionado al tag HA mediante el sistema *UAS-Gal4* que nos permitirá detectar la proteína. Utilizando esta línea analizamos la localización de Haspin en interfase y observamos que se encuentra en el núcleo, acumulándose principalmente en la periferia nuclear, como detectamos mediante la inmunotinción de núcleos de glándulas salivares con el anticuerpo de HA en la línea *UAS-Haspin-HA* activada bajo el promotor del gen *nubbin*. Confirmamos que esta señal es específica ya que no se ve utilizando la línea que sobreexpresa Haspin sin el tag HA (Figura 47).



**Figura 47. Haspin se localiza en la periferia nuclear.** Immunolocalización de HA (verde) en glándulas salivares de larvas *nub>Haspin-HA* (panel superior) y *nub>Haspin* (panel inferior). DNA está teñido con DAPI (azul).

Ya que Haspin está presente en los núcleos cuyo DNA está en interfase nos preguntamos si se une al DNA en esta fase del ciclo. Realizamos una tinción de cromosomas politénicos con HA en esta misma línea, pero no vemos la proteína.

Está descrito que Haspin se une al DNA durante mitosis en otros organismos (Yamagishi et al., 2010) así que nos planteamos si en esta fase del ciclo la podemos ver. Hacemos tinción en cromosomas mitóticos a través de *squash* de células de cerebro. Para ello, activamos la línea con un *driver* general como es *69B-Gal4*, pero tampoco somos capaces de ver en este momento del ciclo su unión al DNA, a pesar de detectar la presencia de la H3T3ph.

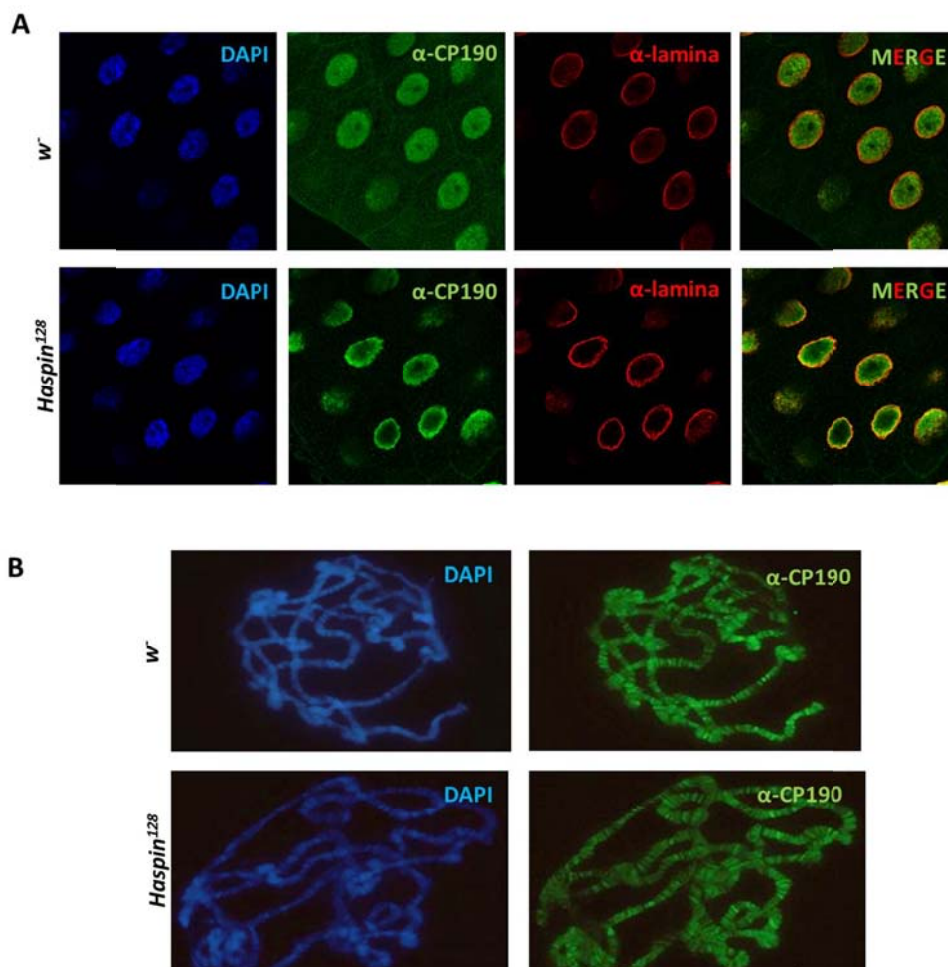
### 3.6.2. La localización de CP190 y la estructura nuclear están alteradas en ausencia de Haspin

Puesto que CP190 es una proteína común a todos los *insulators* descritos hasta el momento en *Drosophila* (Pai et al., 2004, Negre et al., 2010) estudiamos el comportamiento de esta proteína en ausencia de Haspin.

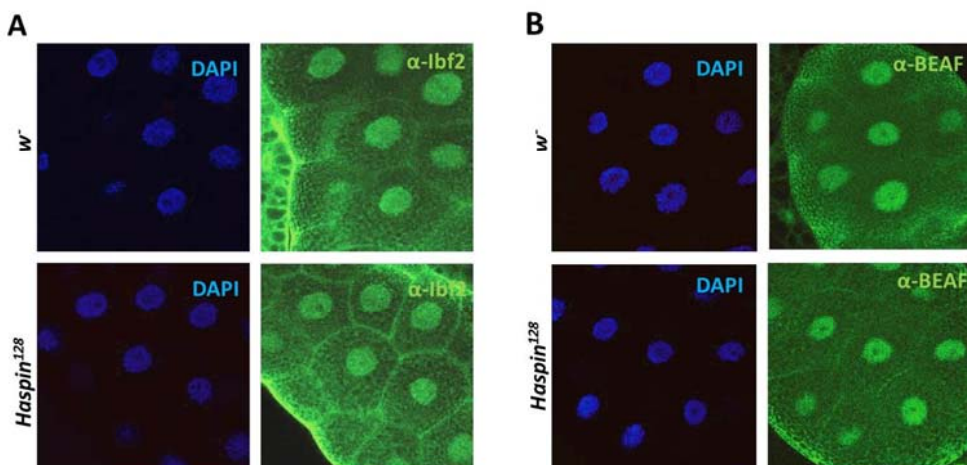
Analizamos mediante inmunotinción la localización de CP190 en núcleos de glándulas salivares. En los individuos control (*w*) se ve que se encuentra en el núcleo. La ausencia de Haspin provoca que tienda a colocarse principalmente en la periferia nuclear (Figura 48A), como se aprecia por una localización más próxima a la lamina nuclear. A través de inmunotinción en cromosomas politénicos no vemos ningún cambio significativo en la asociación de CP190 con el DNA en presencia o ausencia de Haspin (Figura 48B). El efecto sobre CP190 no se observa en otras proteínas con actividad *insulator* tal como Ibf2 (Figura 49A) o BEAF (Figura 49B), ya que éstas siguen manteniendo la misma localización en los núcleos interfásicos en presencia (Figura 49, panel superior) o ausencia de Haspin (Figura 49, panel inferior).

Podemos concluir que en ausencia de Haspin, CP190 continúa asociada a la cromatina pero localiza mayoritariamente en la periferia nuclear.

## Resultados

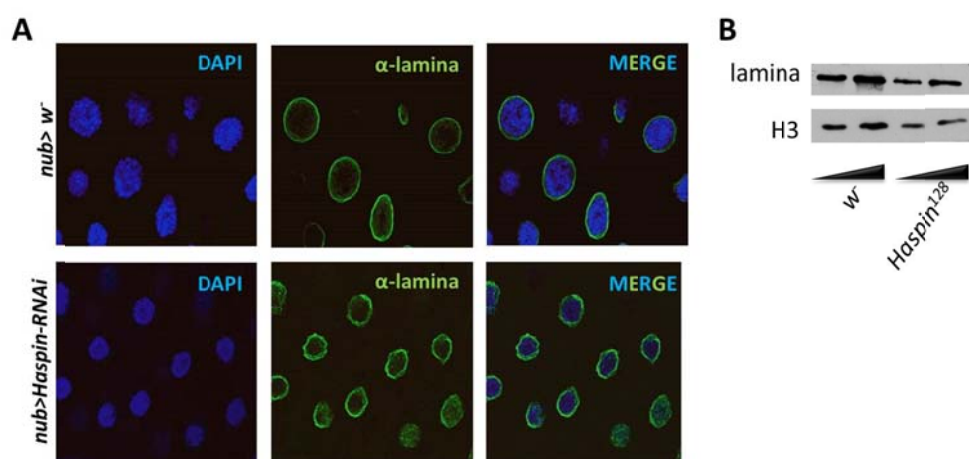


**Figura 48. Haspin afecta a la localización nuclear de CP190 sin interferir con su unión al DNA.** A/ Immunolocalización de CP190 (verde) y lamina (rojo) en núcleos de glándulas salivares de larvas control (*w*<sup>-</sup>) (panel superior) y *Haspin*<sup>128</sup> (panel inferior). DNA está teñido con DAPI (azul). B/ Immunolocalización de CP190 (verde) en cromosomas politénicos en larvas control (*w*<sup>-</sup>) (panel superior) y *Haspin*<sup>128</sup> (panel inferior). DNA está teñido con DAPI (azul).



**Figura 49. Haspin no afecta a la localización nuclear de otras proteínas *insulators*.** Immunolocalización de Ibf2 (A) y BEAF (B) (verde) en núcleos de glándulas salivares de larvas control (*w*<sup>-</sup>) (panel superior) y *Haspin*<sup>128</sup> (panel inferior). DNA está teñido con DAPI (azul).

En los experimentos de inmunolocalización de CP190 en núcleos de glándulas salivares que hemos descrito (Figura 48A) se observa que la lamina se encuentra alterada en ausencia de Haspin. Caracterizamos este fenotipo con mayor profundidad y para ello realizamos inmunolocalización de lamina disminuyendo los niveles de Haspin bajo control del promotor del gen *nubbin*. Como en la línea mutante nula, apreciamos fuertes alteraciones morfológicas, con grandes invaginaciones, otorgándole un aspecto como arrugado (Figura 50A). Esto no se debe a una alteración en los niveles de expresión de la proteína lamina ya que comprobamos mediante ensayos de *Western blot* con extractos totales de proteína de glándulas salivares que la cantidad de lamina es similar en presencia o ausencia de Haspin (Figura 50B).



**Figura 50. Haspin afecta a la estructura nuclear.** A/ Inmunolocalización de lamina (verde) en núcleos de glándulas salivares de larvas control (*nub>w*) (panel superior) y *nub>Haspin-RNAi* (panel inferior). DNA está teñido con DAPI (azul). B/ *Western blot* de extractos de glándulas salivares de larvas control (*w*) y *Haspin*<sup>128</sup> en cantidades crecientes, revelado contra lamina (panel superior) y H3 como control de carga (panel inferior).

En su conjunto, ambos resultados sugieren que Haspin juega un papel en la modulación de la estructura nuclear.

### 3.6.3. La ausencia de Haspin afecta a la condensación de la cromatina

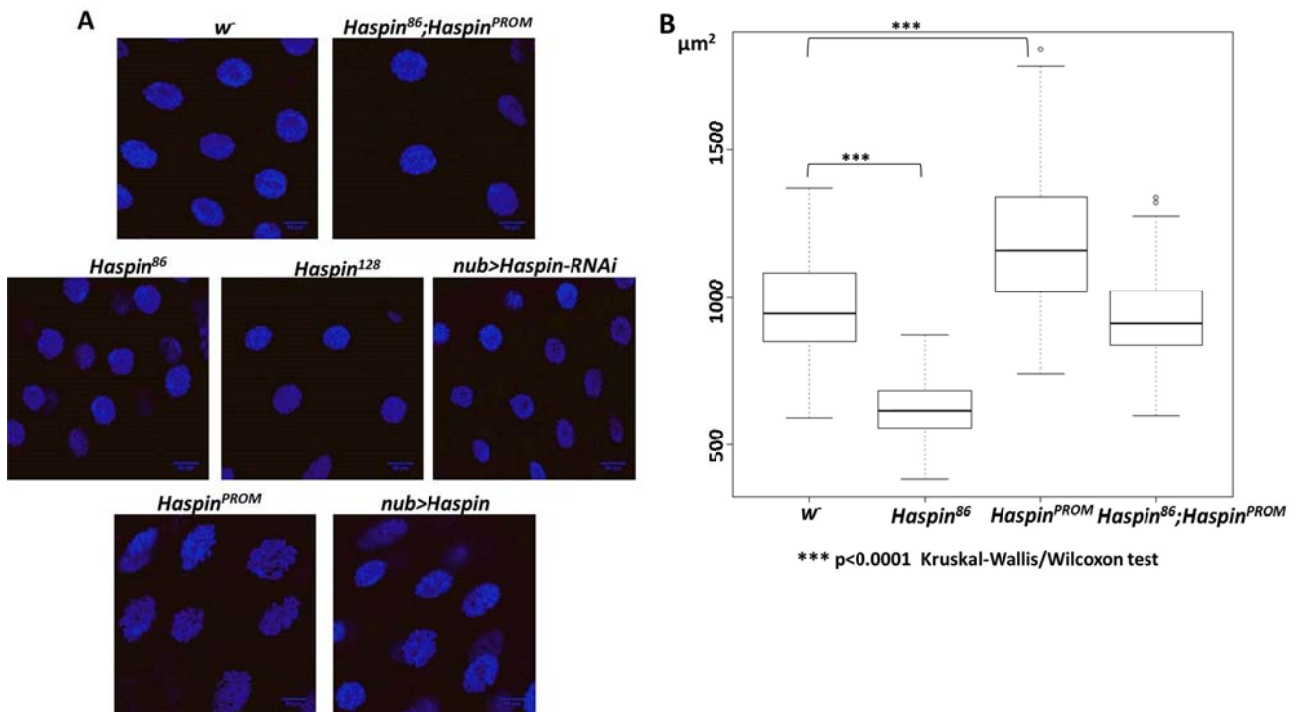
#### 3.6.3.1. Haspin provoca cambios en el tamaño de los núcleos en interfase

En todos los experimentos que hemos realizado de inmunolocalización en núcleos de glándulas salivares hemos detectado que la línea mutante de *Haspin* tiene un tamaño nuclear menor que los control. Para comprobar que esta apreciación es cierta medimos el área de los núcleos de este tejido. El tamaño nuclear lo delimitamos por la señal emitida por DAPI, que tiñe el DNA. Observamos que en ausencia de Haspin el tamaño es menor que en los individuos control, tanto utilizando las dos líneas mutantes como disminuyendo el nivel de expresión de *Haspin* en estas células mediante técnicas de

## Resultados

RNAi. Además, observamos que la sobreexpresión de la proteína, tanto bajo control del promotor propio como por el sistema *UAS-Gal4*, incrementa el tamaño de los núcleos respecto a los control. Comprobamos que este fenotipo se debe a Haspin mediante un experimento de rescate, ya que al expresar Haspin en un fondo mutante para la proteína no se observan diferencias con el control (Figura 51A).

Cuantificamos estos fenotipos y comprobamos que las diferencias entre control y ausencia o sobreexpresión de Haspin son estadísticamente significativas según el test de Kruskal-Wallis/Wilcoxon ( $p < 0.0001$ ). Sin embargo, la diferencia entre control y sobreexpresión de la proteína en fondo mutante no lo es (Figura 51B).

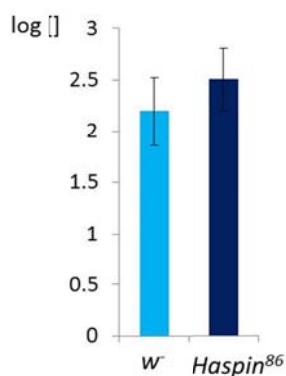


**Figura 51. Haspin afecta al tamaño nuclear.** A/ Immunolocalización de glándulas salivares de larvas control (*w<sup>-</sup>*), de las que sobreexpresan la proteína en fondo mutante (*Haspin<sup>86</sup>;Haspin<sup>PROM</sup>*), de las diferentes mutantes de la proteína (*Haspin<sup>86</sup>*, *Haspin<sup>128</sup>*, *nub>Haspin-RNAi*) y de las que la sobreexpresan (*Haspin<sup>PROM</sup>*, *nub>Haspin*). DNA está marcado con DAPI (azul). B/Gráfico Box plot mostrando la cuantificación del tamaño de los núcleos de glándulas salivares de larvas *w<sup>-</sup>*, *Haspin<sup>86</sup>*, *Haspin<sup>PROM</sup>* y *Haspin<sup>86</sup>;Haspin<sup>PROM</sup>*. \*\*\*  $p < 0.0001$  según test Kruskal-Wallis/Wilcoxon.

### 3.6.3.2. Haspin no afecta a la replicación

Como previamente hemos señalado, los cromosomas politénicos de las glándulas salivares se han formado por sucesivas rondas de replicación del material genético. Por tanto, el cambio en el número de rondas puede afectar al tamaño de núcleos. El menor tamaño que observamos en ausencia de Haspin puede ser debido a que la replicación esté alterada.

Para comprobarlo, extraemos DNA genómico de glándulas salivares, tanto de individuos control (*w*) como de la línea *Haspin*<sup>86</sup> y calculamos mediante PCR cuantitativa la cantidad de DNA en ambas muestras respecto a una recta patrón obtenida con diferentes cantidades conocidas de DNA genómico. El resultado obtenido muestra que la cantidad de DNA no es menor en ausencia de Haspin, más bien al contrario, aunque, según el test t de Student, las diferencias no son estadísticamente significativas ( $p=0.44$ ) (Figura 52). Por tanto, el menor tamaño de los núcleos no es debido a una alteración de la replicación en ausencia de Haspin.



**Figura 52. La mutación de *Haspin* no provoca problemas de replicación del DNA en glándulas salivares.** Representación del logaritmo de la concentración de DNA genómico extraído de glándulas salivares de larvas control (*w*) y *Haspin*<sup>86</sup> midiendo el gen *actina*. Las barras de error muestran la SEM de dos experimentos independientes.

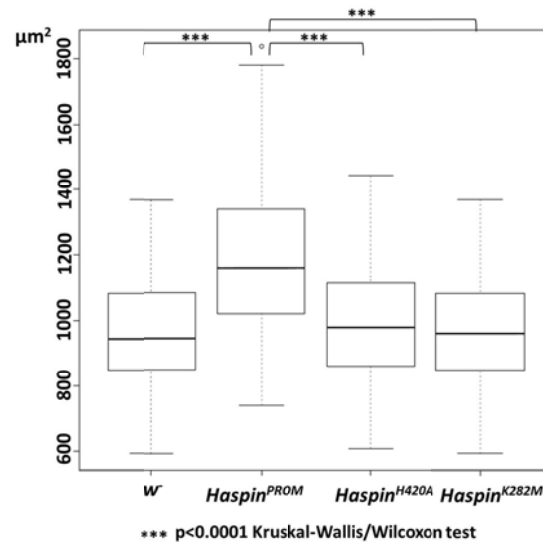
### 3.6.3.3. El fenotipo es producido gracias al dominio kinasa

El dominio kinasa de Haspin es necesario para desempeñar su función en mitosis, ya que la fosforilación de la H3T3 es lo que delimita la localización de Aurora B y el *Cromosomal Passenger Complex* (Yamagishi et al., 2010, Qian et al., 2011, Zhou et al., 2014). Para comprobar si esta parte de la proteína es la que regula el fenotipo que acabamos de describir, hemos obtenido líneas transgénicas que sobreexpresan *Haspin* bajo su propio promotor con su dominio kinasa mutado en la posición 282 en que hemos sustituido la lisina por una metionina, cambio que tradicionalmente se hace para alterar el reconocimiento del ATP en el dominio kinasa. Paralelamente, mutamos la histidina de la posición 420 por una alanina, cambio que se ha descrito para inactivar la kinasa Haspin en diferentes organismos (Eswaran et al., 2009).



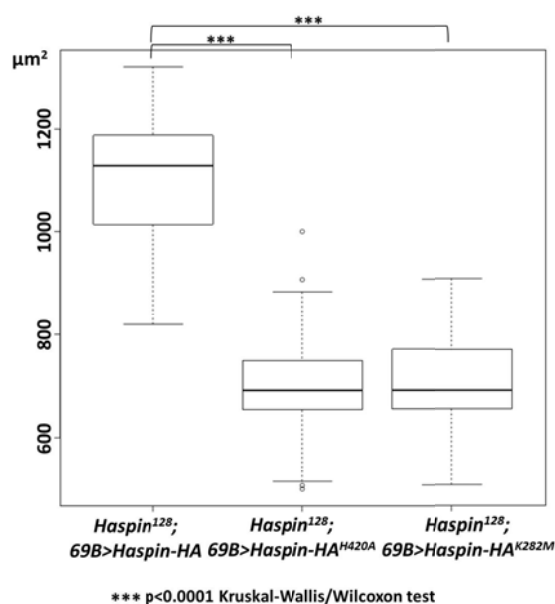
## Resultados

Analizamos en estas líneas transgénicas el tamaño de los núcleos de glándulas salivares tal como hemos hecho previamente. Observamos que mientras que al sobreexpresar Haspin *wild-type* (WT) el tamaño de los núcleos aumenta, no existen diferencias respecto al control (*w*) cuando el dominio kinasa está mutado en cualquiera de las posiciones, según el test de Kruskal-Wallis/Wilcoxon (Figura 53).



**Figura 53. Haspin afecta a la condensación de la cromatina a través de su dominio kinasa.** Gráfico *Box plot* mostrando la cuantificación del tamaño de los núcleos de glándulas salivares en larvas control (*w*), la sobreexpresión de la proteína (*Haspin<sup>PROM</sup>*) y las sobreexpresiones con el dominio kinasa mutado (*Haspin<sup>H420A</sup>* y *Haspin<sup>K282M</sup>*). \*\*\* p<0.0001 según test Kruskal-Wallis/Wilcoxon.

Aunque la única diferencia entre la construcción WT y las mutadas es el cambio de un aminoácido y sabemos que en la línea *Haspin<sup>PROM</sup>* la proteína es funcional, no podemos demostrar que estas líneas expresen la proteína ya que no la detectamos con el anticuerpo. Por ello, creamos unas líneas que sobreexpresan Haspin con el dominio kinasa mutado en las mismas posiciones, unida al tag HA. En este caso la sobreexpresión se realiza bajo el sistema *UAS-Gal4*. Tras comprobar que las líneas realmente sobreexpresan Haspin en glándulas salivares, analizamos el fenotipo sobreexpresando la proteína WT o las mutadas sobre fondo mutante para *Haspin* y observamos que cuando el dominio kinasa está mutado los núcleos son más pequeños (Figura 54). Por tanto, podemos concluir que Haspin regula la condensación nuclear a través de su dominio kinasa.



**Figura 54. Haspin regula la condensación nuclear a través de su dominio kinasa.** Gráfico *Box plot* mostrando la cuantificación del tamaño de los núcleos de glándulas salivares en larvas que sobreexpresan Haspin sobre fondo mutante (*Haspin*<sup>128</sup>;69B>*Haspin*-HA) y dos líneas que lo hacen pero con el dominio kinasa mutado (*Haspin*<sup>128</sup>;69B>*Haspin*-HA<sup>H420A</sup> y *Haspin*<sup>128</sup>; 69B>*Haspin*-HA<sup>K282M</sup>). \*\*\* p<0.0001 según el test Kruskal-Wallis/Wilcoxon.

### 3.7. Análisis de la localización de H3T3ph en el genoma

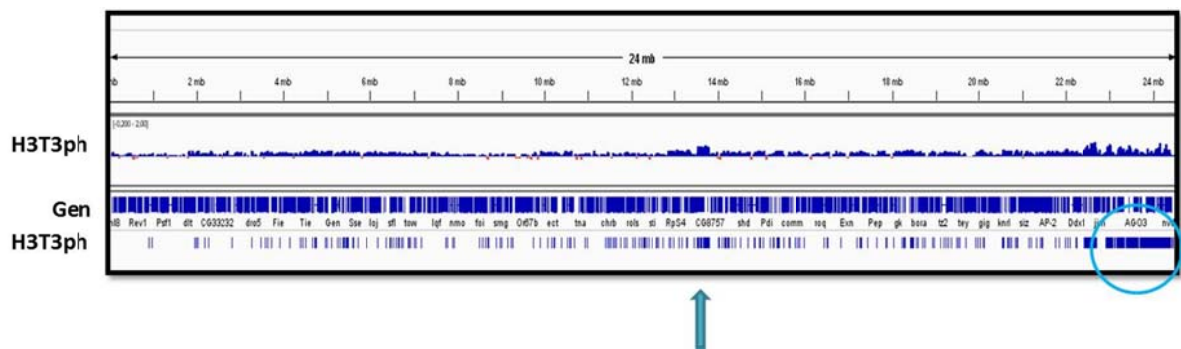
Nuestros resultados muestran que la actividad kinasa de Haspin es necesaria para la regulación de la estructura de la cromatina en interfase y que Haspin fosforila H3T3 en *Drosophila* tanto en mitosis como en interfase, así que caracterizamos la localización de H3T3ph a lo largo de todo el genoma.

Hacemos una inmunoprecipitación de la cromatina con anticuerpo específico de H3T3ph seguida de secuenciación del DNA inmunoprecipitado (ChIP-seq) en células S2. Este experimento permite conocer con suma precisión los sitios a los que una proteína se une al genoma. En el laboratorio previamente hemos hecho ChIP-seq para las proteínas *insulators* CP190 e Ibf2, de manera que disponemos de estos datos para poder compararlos. Además, también hay datos disponibles en modENCODE de diferentes modificaciones de histonas y de otras proteínas de *Drosophila*.

Cuando estudiamos su localización a lo largo de todo un cromosoma, vemos que H3T3ph se acumula preferentemente en regiones pericentroméricas, pero también se observa localizada a lo largo de todos los cromosomas (Figura 57).

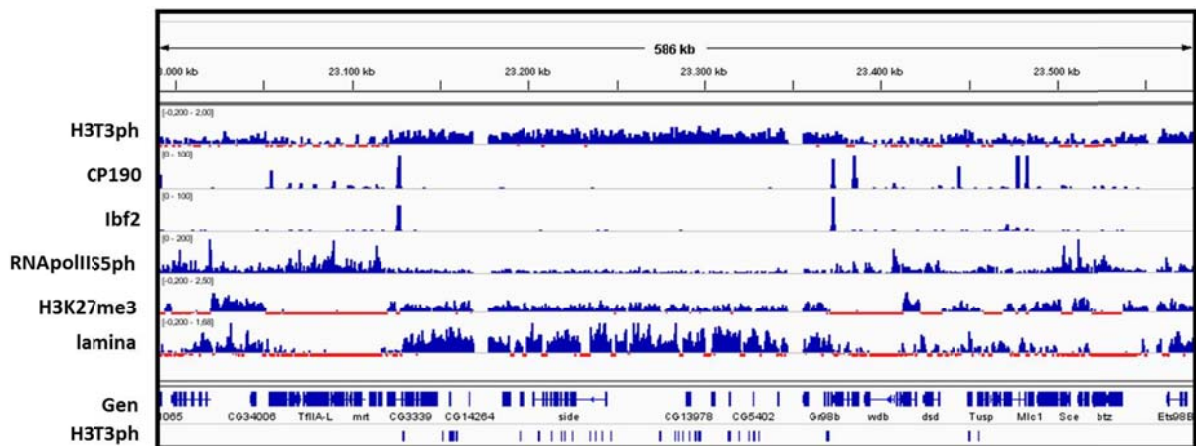


## Resultados



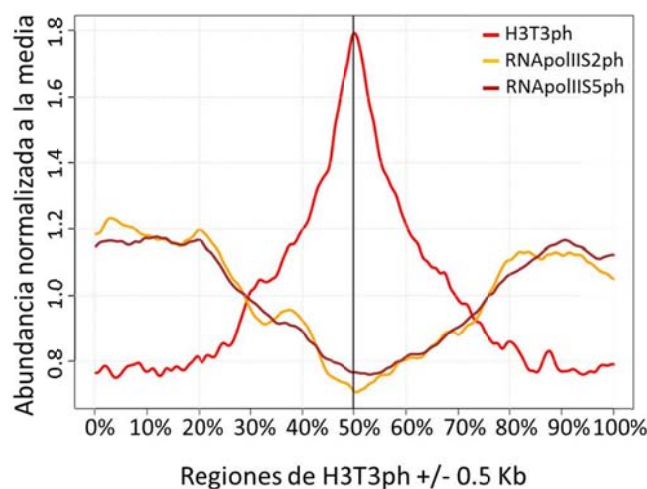
**Figura 55.** H3T3ph se localiza preferente, pero no exclusivamente, en regiones pericentroméricas. Representación del perfil de unión de H3T3ph a lo largo del cromosoma 3L, así como la presencia de zonas enriquecidas. El círculo señala la acumulación en zonas pericentroméricas, y la flecha otra zona también enriquecida.

Si analizamos las zonas donde H3T3ph está presente observamos que no se une a localizaciones discretas como otras proteínas *insulators* tal que CP190 o Ibf2, sino que H3T3ph está en regiones muy amplias, en muchos casos flanqueadas por estas proteínas (Figura 56).



**Figura 56.** Comparación de la unión de H3T3ph con otras proteínas. Representación de una zona enriquecida en H3T3ph en el cromosoma 3R comparándola con el perfil de unión de CP190, Ibf2, RNAPIIIS5ph, H3K27me3 y lamina. En la parte inferior se representan las zonas enriquecidas en H3T3ph.

Respecto a proteínas que se unen a zonas activas del genoma, como puede ser la RNAPIII, apreciamos que las zonas ricas en H3T3ph están libres de esta proteína (Figura 56), y que las pobres en esta marca son las que tienen mayor abundancia en esta proteína. De hecho, si comparamos el perfil de unión a lo largo de todo el genoma de H3T3ph y RNAPIIIS2ph y RNAPIIIS5ph vemos que esta observación se mantiene en todo el genoma, sugiriendo que H3T3ph no está presente en zonas activas (Figura 57).

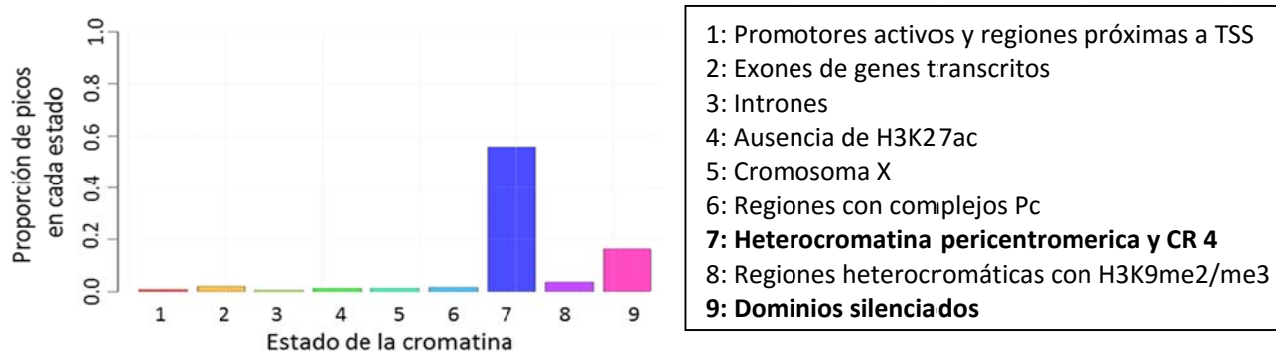


**Figura 57. H3T3ph está ausente en zonas activas.** Representación del perfil de unión de H3T3ph (rojo), RNApolIIIS2ph (amarillo) y RNApolIIIS5ph (morado) a lo largo del genoma en función de la abundancia de H3T3ph en las regiones donde está unida.

Contrastando la unión de H3T3ph con la abundancia de H3K27me3, marca asociada a zonas silenciadas por complejos Polycomb, o con la lamina, unida a zonas silenciadas (Verboon et al., 2015), observamos el efecto opuesto al que acabamos de describir. Las zonas que están reprimidas son las que a su vez están enriquecidas en H3T3ph (Figura 56).

Recientemente se ha hecho una clasificación de la cromatina de *D. melanogaster* en la que se establecen 9 estados en función de las características que tienen en común, tanto de modificaciones de histonas como unión de determinadas proteínas, revelando patrones comunes entre cromosomas, genes, elementos reguladores y otros dominios funcionales (Kharchenko et al., 2011). Esta clasificación provee una visión sobre cómo los genes son regulados, y puede ser una ayuda para investigaciones sobre no solo la estructura del genoma sino también de su función.

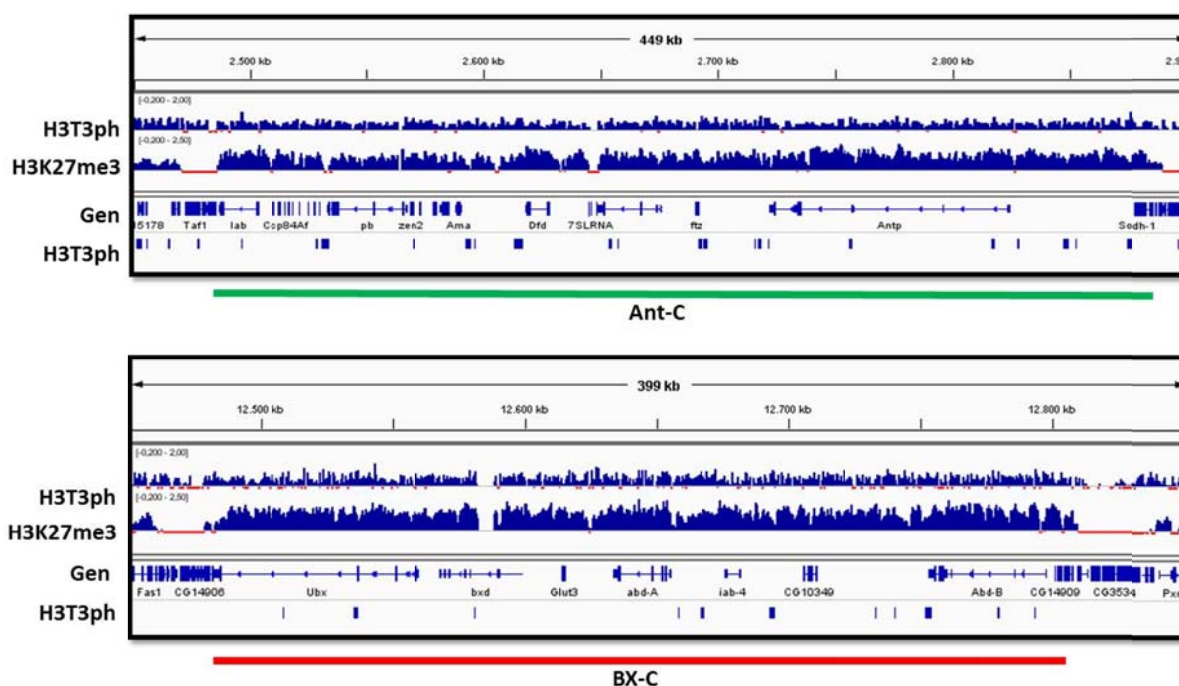
Hemos analizado a qué tipo de cromatina corresponden las zonas enriquecidas en H3T3ph (Figura 58), y observamos que mayoritariamente son heterocromatina pericentromérica y dominios silenciados.



**Figura 58. H3T3ph se une a la heterocromatina pericentromérica y en las regiones silenciadas.** Representación de la unión de H3T3ph a lo largo del genoma en función del tipo de cromatina al que está unida.

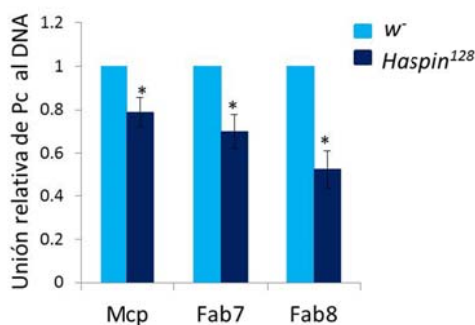
## Resultados

Se sabe que en células S2 los genes homeóticos están silenciados. Estudiamos la unión de H3T3ph a estas regiones y vemos que efectivamente está unida tanto al complejo Antennapedia como al Bithorax (Figura 59).



**Figura 59. H3T3ph se une a genes homeóticos.** Representación del perfil de unión de H3T3ph y H3K27me3 en la región de los complejos Antennapedia (panel superior) y Bithorax (panel inferior). En la parte inferior se representan las zonas enriquecidas en H3T3ph. La barra verde señala el complejo Antennapedia y la roja el complejo Bithorax.

Para comprobar si es la H3T3ph quien regula la unión de Polycomb a los genes silenciados del BX-C, hacemos un CHIP para comparar la unión de Pc en tres localizaciones diferentes del complejo, Mcp, Fab7 y Fab8. Para ello, realizamos la técnica con cerebros, donde se expresa *Abd-B* aunque sus niveles son bajos, de una línea control *w* y de *Haspin*<sup>128</sup>. Al comparar los resultados obtenidos en 4 réplicas vemos que al no estar presente Haspin, la unión de Pc a las zonas analizadas disminuye de manera significativa de acuerdo al t test de Student ( $p < 0.05$ ) (Figura 60).



**Figura 60. Polycomb disminuye su unión a los Fabs del complejo Bithorax en ausencia de Haspin.** Representación relativa de la unión de Pc a Mcp, Fab7 y Fab8 en cerebros de larvas control (*w*) y *Haspin*<sup>128</sup>, respecto a las control. \*  $p < 0.05$  según test t de Student.

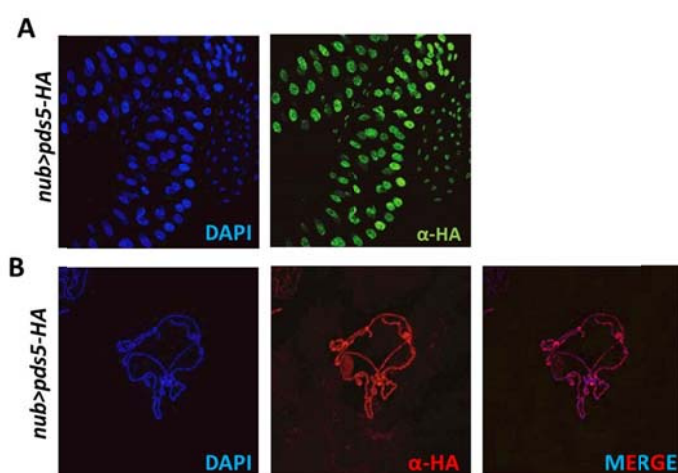
### **3.8. Relación de Haspin con otras proteínas involucradas en la organización de la cromatina**

#### **3.8.1. Cohesinas**

Es conocida la relación que tienen pds5, el complejo de las cohesinas y Haspin durante mitosis (Dai et al., 2006, Yamagishi et al., 2010, Carretero et al., 2013). Analizamos si esta dependencia también se mantiene durante interfase. Para ello creamos una línea transgénica que sobreexpresa pds5 unida al tag HA mediante el sistema *UAS-Gal4*, que nos permite detectarla en ausencia de anticuerpo específico contra la proteína.

##### **3.8.1.1. La unión de pds5 al DNA depende de Haspin**

Sobreexpresamos la proteína pds5-HA en glándulas salivares bajo el promotor del gen *nubbin*. Mediante inmunotinción en este tejido concretamos que es una proteína de localización nuclear (Figura 61A) y por inmunolocalización en cromosomas politénicos vemos que se une al DNA (Figura 61B).

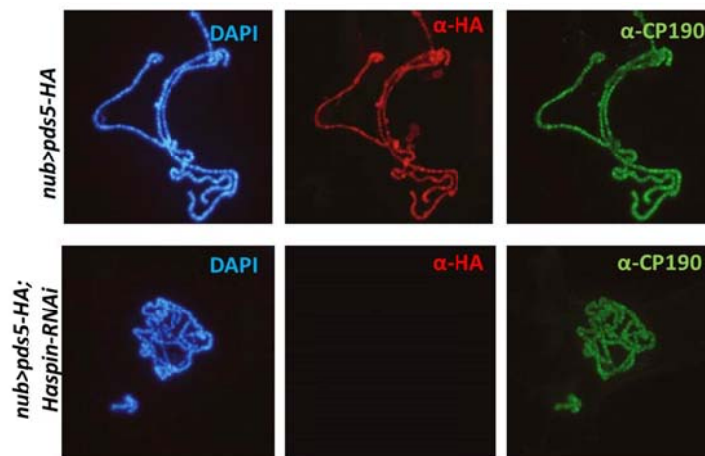


**Figura 61. Pds5 es una proteína nuclear que se une al DNA.** A/ Immunolocalización de pds5-HA (verde) en núcleos de glándulas salivares de larvas *nub>pds5-HA*. DNA está teñido con DAPI (azul). B/ Immunolocalización de pds5-HA (rojo) en cromosomas politénicos de larvas *nub>pds5-HA*. DNA está teñido con DAPI (azul).

Se ha descrito una interdependencia para la localización de pds5-cohesinas-Haspin en mitosis en otros organismos (Dai et al., 2006, Yamagishi et al., 2010). Analizamos qué ocurre en *Drosophila* durante interfase. Para ello, sobreexpresamos pds5-HA bajo control del promotor del gen *nubbin* en presencia o ausencia de Haspin. Mediante inmunolocalización en cromosomas politénicos, vemos

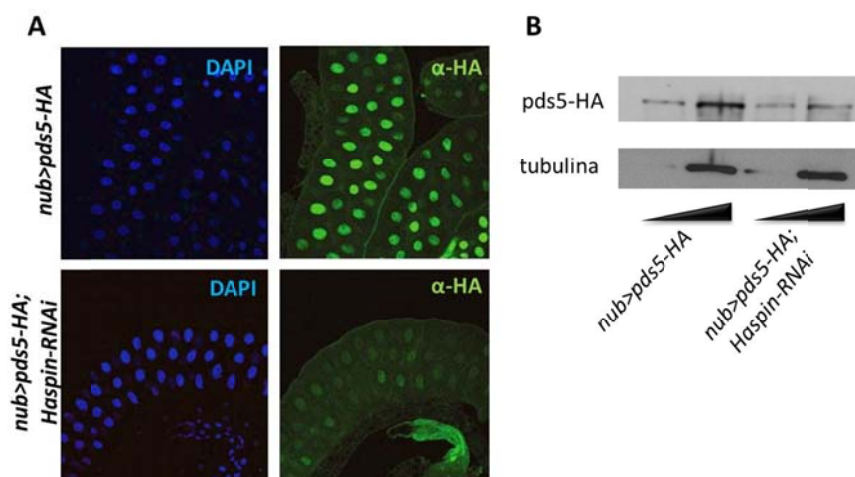
## Resultados

que la ausencia de Haspin provoca que pds5 deje de estar unida al DNA. Este efecto es específico para pds5, ya que no afecta a la unión de otras proteínas como CP190 (Figura 62).



**Figura 62. Haspin afecta a la unión de pds5 al DNA.** Immunolocalización de pds5-HA (rojo) y CP190 (verde) en cromosomas politénicos de larvas *nub>pds5-HA* (panel superior) y *nub>pds5-HA;Haspin-RNAi* (panel inferior). DNA está teñido con DAPI (azul).

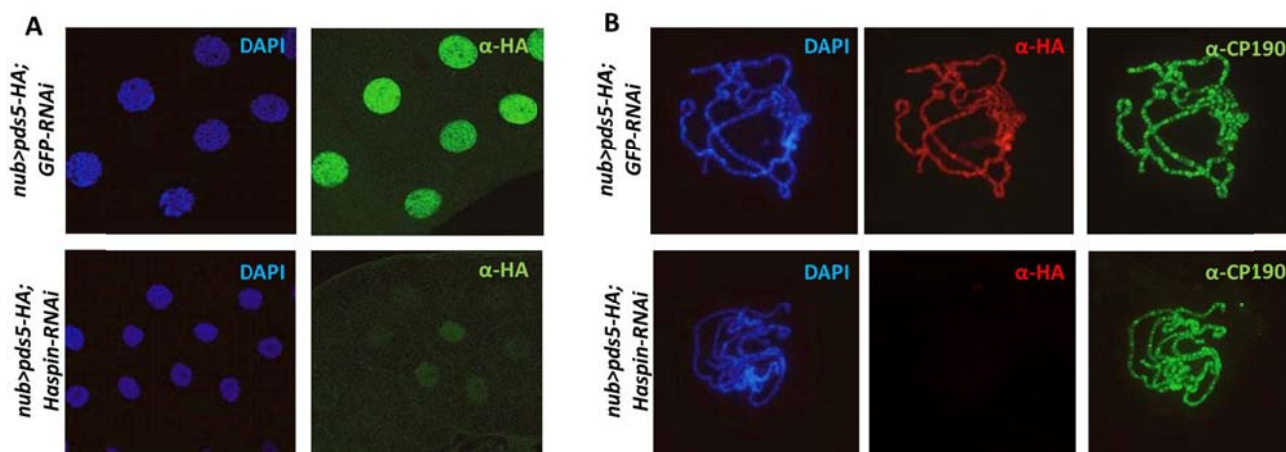
Realizamos immunolocalización de pds5-HA en núcleos de glándulas salivares en presencia o ausencia de Haspin. Pds5 sigue estando presente en el núcleo en ausencia de Haspin, aunque la cantidad de proteína es menor que cuando Haspin está en el medio (Figura 63A). Cuantificamos la cantidad de proteína total de un extracto de glándulas salivares mediante *Western Blot* y observamos que el nivel de proteína total disminuye un 50% en ausencia de Haspin, respecto a la cantidad detectada cuando la proteína está presente en el medio (Figura 63B).



**Figura 63. La cantidad de pds5 es menor en ausencia de Haspin.** A/ Immunolocalización de pds5-HA (verde) en núcleos de glándulas salivares de larvas *nub>pds5-HA* (panel superior) y *nub>pds5-HA;Haspin-RNAi* (panel inferior). DNA está teñido con DAPI (azul). B/ *Western blot* de extractos proteicos de glándulas salivares de larvas *nub>pds5-HA* y *nub>pds5-HA;Haspin-RNAi* en cantidades crecientes, revelado contra HA (panel superior) y tubulina como control de carga (panel inferior).

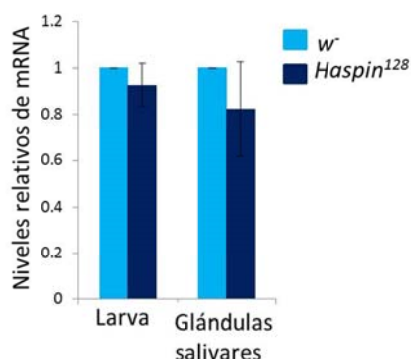


Repetimos el experimento pero utilizando en este caso el sistema *UAS-Gal4* con líneas que sobreexpresan *pds5*-HA a la vez que producen RNAi contra *GFP*, que nos sirve como control, o contra *Haspin*, para en ambos casos tener *Gal4* activando RNAi y la sobreexpresión. Tanto mediante la inmunolocalización de *pds5*-HA en núcleos de glándulas salivares (Figura 64A) como en cromosomas politénicos (Figura 64B) seguimos observando el mismo resultado. De esta manera confirmamos que la ausencia de *Haspin* afecta a la cantidad total de *pds5* y a su capacidad de unión al DNA.



**Figura 64. La cantidad de *pds5* y su unión al DNA dependen de *Haspin*.** A/ Inmunolocalización de *pds5*-HA (verde) en núcleos de glándulas salivares de larvas *nub>pds5-HA;GFP-RNAi* (panel superior) y *nub>pds5-HA;Haspin-RNAi* (panel inferior). DNA está teñido con DAPI (azul). B/ Inmunolocalización de *pds5*-HA (rojo) y CP190 (verde) en cromosomas politénicos de larvas *nub>pds5-HA;GFP-RNAi* (panel superior) y *nub>pds5-HA;Haspin-RNAi* (panel inferior). DNA está teñido con DAPI (azul).

Para caracterizar si la disminución en la cantidad de *pds5* es debida a regulación transcripcional o posttranscripcional, extraemos mRNA tanto de larvas mutantes de *Haspin* como de larvas control, así como de glándulas salivares de ambas líneas. Tras retrotranscribir este RNA, vemos mediante PCR cuantitativa que el nivel de expresión de *pds5* no varía en presencia o ausencia de *Haspin* (Figura 65).



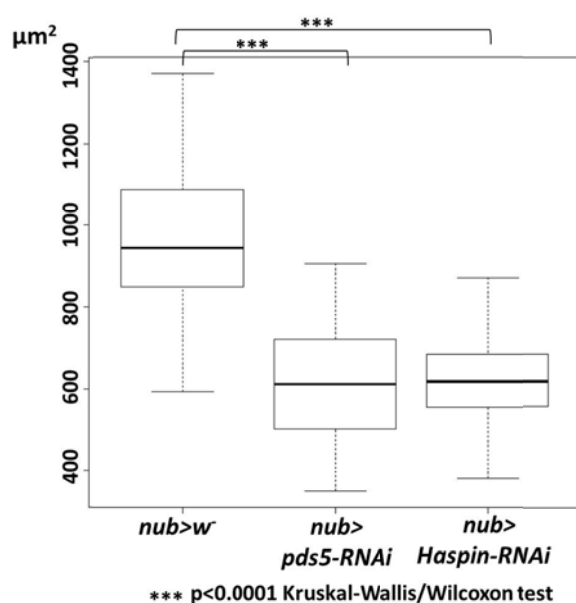
**Figura 65. El nivel de expresión de *pds5* no depende de *Haspin*.** Representación relativa del nivel transcripcional del gen *pds5* en larvas control (*w<sup>+</sup>*) y *Haspin<sup>128</sup>* respecto a las control así como en glándulas salivares. Las barras de error muestran los SEM de tres experimentos independientes.

Por tanto, nuestros resultados sugieren que *pds5* se degrada parcialmente en ausencia de *Haspin*.

## Resultados

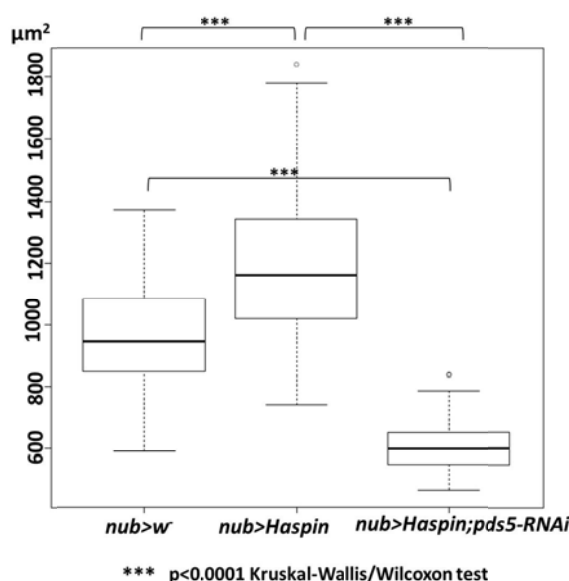
### 3.8.1.2. Pds5 regula la condensación de la cromatina

Dorsett y sus colaboradores mostraron que la ausencia de *pds5* provocaba alteraciones en la condensación de los cromosomas politénicos (Dorsett et al., 2005). Analizamos el tamaño de los núcleos disminuyendo los niveles de *pds5* en glándulas salivares bajo control del promotor del gen *nubbin*. Comprobamos que el tamaño es menor que en los individuos control *nub>w<sup>+</sup>* (Figura 68), de manera estadísticamente significativa según el test de Kruskal-Wallis / Wilcoxon, y similar al tamaño observado al hacer RNAi contra *Haspin* (Figura 66).



**Figura 66. Pds5 regula la condensación de la cromatina.** Gráfico *Box plot* mostrando la cuantificación del tamaño de los núcleos de glándulas salivares en larvas control (*nub>w<sup>+</sup>*), *nub>UAS-pds5-RNAi* y *nub>UAS-Haspin-RNAi*. \*\*\* p<0.0001 según test Kruskal-Wallis/Wilcoxon.

Puede que el efecto producido por ambas líneas tenga la misma causa, que la disminución sea debida a que *Haspin* necesita a *pds5* para unirse al DNA y ejercer su efecto en la condensación, como está descrito para mitosis en otros organismos (Yamagishi et al., 2010). Para estudiar si existe una relación, eliminamos *pds5* en glándulas salivares en una línea que sobreexpresa *Haspin* en este tejido. Observamos que a pesar de sobreexpresar *Haspin* no se produce el aumento de tamaño de los núcleos que previamente habíamos detectado, y que *pds5* tiene que estar presente para ver este fenotipo (Figura 67). Este resultado sugiere que en nuestro animal modelo existe una interdependencia entre *pds5* y *Haspin* para realizar correctamente su función en interfase.

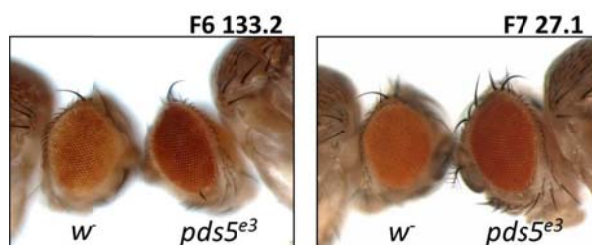


**Figura 67. Haspin y pds5 regulan la condensación nuclear.** Gráfico *Box plot* mostrando la cuantificación del tamaño de los núcleos de glándulas salivares en larvas control (*nub>w*), *nub>Haspin* y *nub>Haspin;pds5-RNAi*. \*\*\* p<0.0001 según test t Kruskal-Wallis/Wilcoxon.

### 3.8.1.3. Pds5 y cohesinas tienen actividad *enhancer-blocking*

Estudiamos mediante ensayos de *enhancer-blocking* si pds5 y una de las proteínas del complejo de las cohesinas, SMC1, pueden tener también esta actividad en *Drosophila*.

Llevamos a cabo este ensayo con varias líneas transgénicas con *insulators* tanto del Fab6 (F6 133.2 y F6 26.3C) como del Fab7 (F7 B+P 2.5 y F7 27.1) entre el *enhancer* y el promotor del gen *white*. De los descendientes obtenidos de los cruces genéticos de cada una de estas líneas con mutantes nulos de *pds5*, *pds5<sup>e3</sup>* (Dorsett et al., 2005) o con control, *w*, detectamos que en todas las líneas analizadas se produce incremento en la pigmentación de los ojos en ausencia de pds5 respecto a los control, lo que indica que esta proteína participa en la actividad *enhancer-blocking* de estos elementos (Figura 68 muestra el ensayo con las líneas F6 133.2 y F7 27.1).

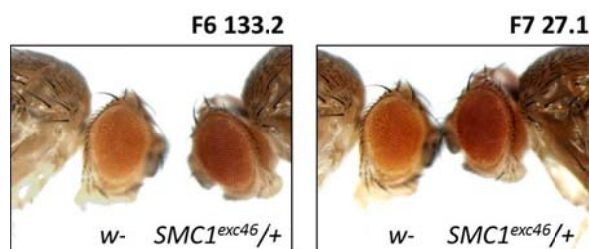


**Figura 68. Pds5 tiene actividad *enhancer-blocking*.** Color de ojos de moscas representativas obtenidas del cruce genético de las líneas transgénicas indicadas en cada imagen con moscas control (*w*) o *pds5<sup>e3</sup>*, presentando la mutación en heterozigosis.



## Resultados

Para valorar la actividad de SMC1 realizamos el ensayo con las mismas líneas transgénicas del Fab6 y Fab7, utilizando una línea con una mutación amórfica para la cohesina SMC1, *SMC1<sup>exc46</sup>*. Con todas las líneas observamos efecto respecto a los control, lo que indica que SMC1 también está jugando un papel en la actividad *enhancer-blocking* (Figura 69 muestra el ensayo con las líneas F6 133.2 y F7 27.1).



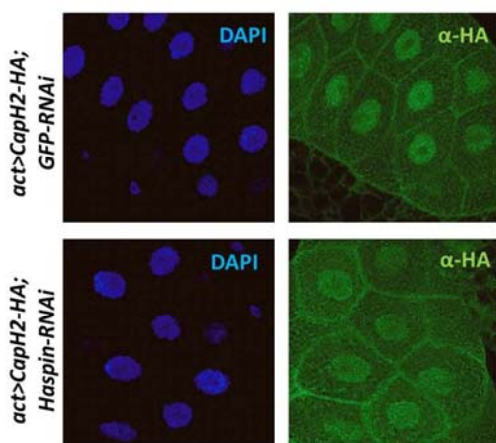
**Figura 69. SMC1 tiene actividad *enhancer-blocking*.** Color de ojos de moscas representativas obtenidas del cruce genético de las líneas transgénicas indicadas en cada imagen con moscas control (*w*) o *SMC1<sup>exc46</sup>*, presentando la mutación en heterocigosis.

### 3.8.2. Condensinas

#### 3.8.2.1. La presencia nuclear de CapH2 no depende de Haspin

Para estudiar si existe alguna relación entre Haspin y el complejo de las condensinas II, creamos una línea transgénica que sobreexpresa CapH2, miembro de este complejo, unida al tag HA, que nos permite detectarla mediante el sistema *UAS-Gal4* en ausencia de anticuerpo específico contra la proteína.

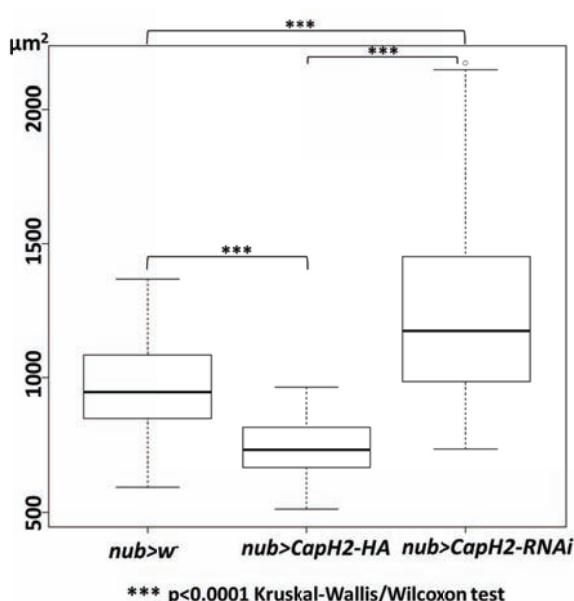
Al sobreexpresar CapH2-HA bajo el promotor de *actina* observamos mediante inmunotinción de glándulas salivares que la proteína está presente en los núcleos en interfase (Figura 70, panel superior), y que la ausencia de Haspin no altera su localización (Figura 70, panel inferior).



**Figura 70. La presencia nuclear de CapH2 no depende de Haspin.** Inmunolocalización de CapH2-HA (verde) en núcleos de glándulas salivares de larvas control (*act>CapH2-HA;GFP-RNAi*) (panel superior) y *act>CapH2-HA;Haspin-RNAi* (panel inferior). DNA está teñido con DAPI (azul).

### 3.8.2.2. Los efectos de Haspin sobre la condensación de la cromatina dependen de la presencia de CapH2

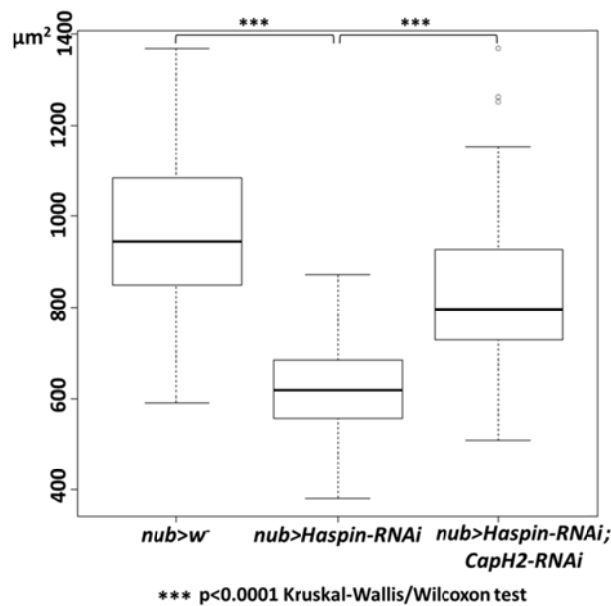
La sobreexpresión de las condensinas provoca condensación del material genético (Buster et al., 2013), como comprobamos al cuantificar el tamaño de los núcleos de glándulas salivares en una línea que sobreexpresa CapH2 bajo el promotor del gen *nubbin* (Figura 71). El efecto contrario sucede cuando esta proteína está ausente; eliminamos CapH2 bajo el promotor del gen *nubbin*, y al cuantificar el tamaño de los núcleos en el mismo tejido, comprobamos que se produce un aumento del mismo respecto de los control *nub>w* (Figura 71).



**Figura 71. CapH2 regula la condensación de la cromatina.** Gráfico *Box plot* mostrando la cuantificación del tamaño de los núcleos de glándulas salivares en larvas control (*nub>w*), *nub>CapH2-HA* y *nub>CapH2-RNAi*. \*\*\* p<0.0001 según test Kruskal-Wallis / Wilcoxon.

Hemos demostrado que la ausencia de Haspin provoca condensación de los núcleos interfásicos (apartado 3.6.3.). Analizamos el tamaño de los núcleos cuando eliminamos ambas proteínas, Haspin y CapH2, y observamos que el fenotipo producido por la falta de Haspin se recupera en parte (Figura 72). Este resultado sugiere que la condensación producida en ausencia de Haspin podría deberse a una unión excesiva del complejo de las condensinas II a la cromatina.

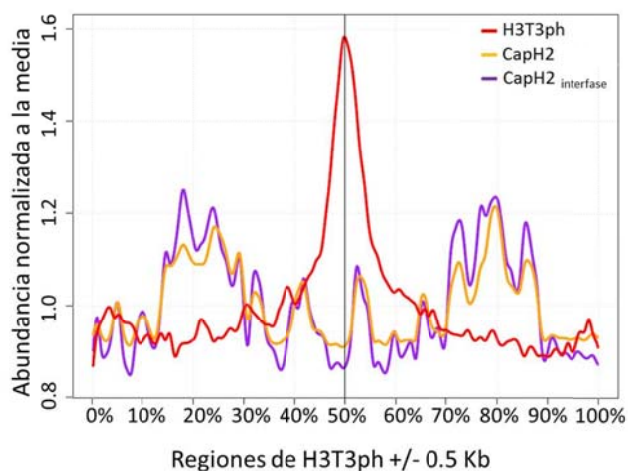
## Resultados



**Figura 72.** El efecto en la condensación que produce la falta de Haspin requiere de la presencia de CapH2. Gráfico *Box plot* mostrando la cuantificación del tamaño de los núcleos de glándulas salivares en larvas control (*nub>w*), *nub>Haspin-RNAi* y *nub>Haspin-RNAi;CapH2-RNAi*. \*\*\* p<0.0001 según test Kruskal-Wallis / Wilcoxon.

### 3.8.2.3. Los sitios enriquecidos en H3T3ph no tienen CapH2, y viceversa

CapH2 es una proteína de unión al DNA (Smith et al., 2013). Cuando comparamos el perfil de unión a la cromatina de H3T3ph respecto al descrito para CapH2, vemos que en las regiones enriquecidas en la modificación que produce Haspin, CapH2 no está presente, y al contrario, las regiones donde la condensina muestra la mayor afinidad son pobres en H3T3ph (Figura 73). Esto sugiere que la presencia de H3T3ph pueda regular negativamente la unión de CapH2, por lo que las zonas ricas en la marca de la histona no tendrían la condensina, y, donde no está presente, CapH2 sería capaz de unirse.



**Figura 73. Perfil de unión de H3T3ph y CapH2 a la cromatina.** Representación del perfil de unión de H3T3ph (rojo), CapH2 (amarillo) y CapH2 en interfase (morado) a lo largo del genoma en función de la abundancia de H3T3ph en las regiones donde está unida.

# Discusión



En el laboratorio estamos interesados en los factores y mecanismos involucrados en la organización de la cromatina en diferentes dominios y su participación en la regulación de la expresión génica.

Durante el desarrollo de esta tesis hemos abordado este estudio desde dos ópticas diferentes; por un lado, caracterizando nuevas funciones de proteínas *insulators* o arquitectónicas que son las proteínas que separan estos dominios, y por otro lado identificando nuevas proteínas con actividad organizadora de la cromatina. En el primer apartado hemos caracterizado la participación de CTCF en la regulación de la síntesis de hormonas esteroideas en *Drosophila* y en el segundo hemos analizado la participación de las proteínas Ibf1, Ibf2 y Haspin en la organización de la cromatina dentro del núcleo en *Drosophila*.

## **1. NUEVAS FUNCIONES DE PROTEÍNAS *INSULATORS***

### **1.1. CTCF regula el tiempo de desarrollo en *Drosophila***

Las proteínas con actividad *insulator*, y en particular CTCF, juegan un papel importante en la organización de la cromatina así como en la expresión de genes durante el desarrollo, tanto en mamíferos como en *Drosophila* (Phillips and Corces, 2009, Herold et al., 2012). Probablemente *Drosophila* sea el animal modelo más apropiado para el estudio de los mecanismos moleculares que regulan el desarrollo debido a sus drásticos cambios entre los estadios larvarios y la metamorfosis. Estas transiciones están bajo un estricto control endocrino, en el que diversas hormonas y neuropéptidos están involucrados. Sin duda, la hormona esteroidea ecdisona es el regulador principal. En los últimos años la aplicación de técnicas de genética molecular de *Drosophila* ha dado nuevas perspectivas así como estrategias para el estudio de la señalización de ecdisona (Gilbert et al., 2002) .

La síntesis de ecdisona está regulada por varios factores, algunos de ellos externos a la glándula protorácica, donde ésta se produce, como la influencia ejercida por la hormona protoracicotrópica (PTTH) o la insulina, o bien de dentro de la propia glándula, como los genes *Halloween*. Además, puede que no esté alterada la síntesis sino su conversión en la forma activa o la respuesta de los tejidos a la misma, procesos que tienen lugar una vez liberada la hormona de la PG.

El principal neuropéptido responsable de la producción de ecdisona es PTTH, y se le ha considerado el factor fundamental que determina el momento de las transiciones del desarrollo. Su falta retrasa el comienzo de la metamorfosis, pero algunas larvas consiguen pupar, por lo que no es necesario para la metamorfosis, pero sí para regular el tiempo en el que se produce (McBrayer et al., 2007). No se conoce qué es lo que regula la función de PTTH. Se ha visto que la expresión del gen fluctúa enormemente durante el desarrollo (Mirth et al., 2005) y que algunos factores que le afectan (como los ciclos de luz/oscuridad) producen cambios en los niveles transcripcionales del gen. A pesar

## **Discusión**

de que en ausencia de PTTH o CTCF los fenotipos observados sean los mismos (retraso en el desarrollo pero finalmente se realiza la pupación) hemos detectado que en ausencia de CTCF los niveles de expresión de *PTTH* no disminuyen, y que la expresión de su receptor (*Torso*) (Rewitz et al., 2009) es independiente de CTCF. De hecho, cuando quitamos específicamente CTCF en la glándula protorácica la expresión de *PTTH* incluso aumenta. Estudios fisiológicos confirman la existencia de *feedbacks* tanto positivos como negativos con influencia en la síntesis de ecdisona en la PG. Si todos los cambios que se producen tras la esteroidogénesis están predeterminados, es razonable pensar que la producción y liberación de PTTH está influenciada por este tipo de control. No obstante, esta posibilidad no ha sido estudiada a nivel molecular, ya que en ninguno de estos *feedbacks* se ha involucrado a PTTH como factor que participe en el proceso (Yamanaka et al., 2013), por lo que no podemos confirmar que este aumento de PTTH en ausencia de CTCF sea consecuencia de uno de estos mecanismos.

PTTH es el principal estímulo para inducir la producción de ecdisona, pero no el único; muchos factores, sobre todo nutricionales que reflejan el estado metabólico general de la larva, actúan directamente sobre la glándula protorácica para controlar la síntesis de ecdisona y, en consecuencia, el desarrollo. Se ha confirmado la participación de las vías de insulina así como la de *Target of Rapamycin* (TOR), que tienen un papel directo en la regulación de la síntesis de ecdisona en la PG (Caldwell et al., 2005, Colombani et al., 2005, Mirth et al., 2005, Layalle et al., 2008). Estudios previos ya habían relacionado las señales mediadas por la insulina como un regulador de crecimiento (Oldham and Hafen, 2003, Hietakangas and Cohen, 2009) y lo habían asociado con la activación de varios factores que actuaban posteriormente en la PG como PI3K y AKT. Se ha visto que CTCF juega un papel en la proliferación celular inducida por insulina y actúa tras la activación por la hormona de Erk y AKT (Gao et al., 2007, Tsui et al., 2014).

Por otro lado, estudios recientes indican que la señalización vía TGF $\beta$ /activin es también necesaria para la síntesis de ecdisona y las transiciones en el desarrollo (Gibbens et al., 2011). El principal mediador que actúa bajo la influencia de la señal de TGF $\beta$ , dSmad2, participa directamente en la regulación de la transcripción de los genes *Halloween*. El hecho de que TGF $\beta$  promueva la formación de complejos entre Smad y CTCF en la región reguladora de H19 (Bergstrom et al., 2010), sugiere la posibilidad de que estos complejos también puedan regular la transcripción de los genes *Halloween*.

Los cambios fisiológicos, morfológicos y de comportamiento relacionados con las transiciones larvarias y la metamorfosis son producidos por 20E, forma activa que adquiere la ecdisona una vez ha sido liberada a la hemolinfa, gracias a la enzima shade. La unión de 20E a su receptor activa la expresión de los genes en diferentes tejidos, por lo que las diferentes respuestas que se dan a esta unión pueden deberse a alteraciones en la normal acumulación de la hormona. A pesar de los bajos niveles transcripcionales de algunos de los genes *Halloween* en ausencia de CTCF, algunas larvas consiguen suficiente cantidad de hormona como para hacer la metamorfosis. Aunque no hemos analizado los niveles de la hormona, hemos visto que los niveles transcripcionales de varios genes que responden a ecdisona, como son *E74B* y *E75B*, no dependen de CTCF, aunque estos niveles son

alcanzados dos días después debido al retraso en el desarrollo que produce la falta de la proteína *insulator*. Al evaluar la expresión del receptor de ecdisona que transmite la señal a estos genes no observamos ningún cambio, ni en cantidad ni en tiempo de expresión. Ya que ni los niveles transcripcionales del receptor ni los de los genes que responden a ecdisona se ven alterados, concluimos que CTCF no influye en la respuesta a la ecdisona, pero sí en el tiempo de su síntesis.

Una propiedad que se conserva tanto en vertebrados como en invertebrados durante la esteroidogénesis es la implicación de las enzimas citocromo P450 (CYP) (Niwa and Niwa, 2014). Los genes que codifican estas proteínas en insectos son los genes *Halloween*, cuya mutación conduce a la muerte embrionaria, lo que enfatiza su importancia fisiológica (Chavez et al., 2000, Warren et al., 2002, Petryk et al., 2003, Niwa et al., 2004, Warren et al., 2004). Son los responsables de la conversión de colesterol a la forma activa de la ecdisona. Durante el desarrollo larvario todos se expresan en las células de la glándula protorácica, con la excepción de *shade*, que lo hace en los tejidos diana de la ecdisona. La regulación de la esteroidogénesis se hace a través del control transcripcional de estas enzimas (Keightley et al., 1990, Sewer and Waterman, 2003). Los niveles de expresión de los genes *Halloween* durante el desarrollo larvario en la PG, como hemos señalado previamente para la ecdisona, aumentan al final de cada estadio, disminuyen al comienzo del siguiente, y vuelven a aumentar al final de la transición (McBrayer et al., 2007). Se sabe que están regulados transcripcionalmente pero no cómo. En los últimos años se ha ido avanzando en este campo y se han propuesto como candidatos de este proceso algunos factores de transcripción así como modificadores de la cromatina (Parvy et al., 2005, Pankotai et al., 2010, Deng and Kerppola, 2013, Moeller et al., 2013, Danielsen et al., 2014) pero aún queda mucho por conocer.

Confirmamos que CTCF afecta a su expresión, ya que tanto la ausencia de la proteína en todo el organismo como de manera concreta en la PG produce cambios en la expresión de algunos de estos genes a lo largo del desarrollo. Sin embargo, CTCF no influye en todos ellos de igual manera. Mientras que en algunos tal que *phantom*, *neverland* o *shroud* la ausencia de CTCF no produce cambios significativos, en *disembodied* genera un retraso en el incremento de la expresión y en *spookier*, *shadow* y *noppera-bo* los niveles de expresión alcanzados son claramente inferiores. Las secuencias de todos estos genes, a excepción de *nvd*, tienen sitios de unión para CTCF próximos a su TSS. CTCF colocaliza con las proteínas *insulators* CP190, Ibf1 e Ibf2 en varios sitios del genoma (Cuartero et al., 2014). Estudiamos si éstas tienen algún efecto en este proceso y vemos que el papel desempeñado por CTCF es independiente de ellas. Queda pendiente determinar si la influencia que ejerce CTCF en la expresión de los genes *Halloween* se realiza a través de los sitios de unión que presentan sus secuencias y si complejos con otros factores transcripcionales pudieran estar relacionados.



### **1.2. CTCF es necesaria para la homeostasis del colesterol**

Hemos señalado que tanto PTTH como los genes *Halloween* son necesarios para la síntesis de ecdisona, y por tanto, para que los tejidos diana respondan. Los efectos que causa tanto la carencia del neuropéptido como la falta de las enzimas son revertidos si se añade al medio la forma activa de la hormona, 20E, de manera que no sea necesaria su síntesis, como se ha visto en diferentes estudios (Ono et al., 2006, McBrayer et al., 2007). Esta recuperación del fenotipo confirma que afectan a la vía de producción de la hormona y no a la respuesta a la misma. Realizando este mismo experimento nosotros somos capaces de ver una recuperación parcial pero no total. Se ha descrito que el efecto de suplementar 20E en el tiempo necesario para comenzar la metamorfosis no es tan evidente como lo es en el caso de la ecdisona, cuya adición al medio lleva a un adelantamiento en su inicio, a pesar de producir una posterior muerte en caso de no haber adquirido el peso mínimo suficiente para poder sostener todos los cambios que se producen en este proceso (Ono, 2014). No obstante, nosotros observamos lo mismo añadiendo ecdisona a cuando probamos con 20E, sólo una cierta recuperación (datos no mostrados). Esto sugiere que la influencia de CTCF en la correcta expresión temporal de los genes diana de la ecdisona no es solo sobre la vía de síntesis de la hormona, sino que está implicado en algún otro proceso.

Los insectos no son capaces de sintetizar el colesterol de *ново* desde acetato, por lo que se ven obligados a tomarlo del medio para convertirlo en ecdisona (Clark and Block, 1959). Por tanto, la ausencia del mismo (o de otro esteroide a partir del cual sean capaces de comenzar la biosíntesis) es un factor limitante. Se ha visto que en el momento de la pupación hay un incremento de lípidos en la glándula protorácica (Talamillo et al., 2013). La falta de CTCF produce un incremento de esta acumulación de grasa al final del desarrollo larvario lo que indica que la homeostasis de lípidos en la glándula está alterada en mutantes de CTCF. Con la adición de un exceso de colesterol en el medio conseguimos una recuperación parcial del fenotipo, corroborando que CTCF ejerce esta acción sobre las grasas. Al suplementar la alimentación con 20E y colesterol al mismo tiempo se observa la total recuperación del retraso en el desarrollo, confirmando que CTCF participa tanto en la regulación de la expresión génica de las enzimas involucradas en la síntesis de ecdisona como en la homeostasis de colesterol. ¿De qué manera actúa en este proceso?

La agregación subcelular de lípidos que se produce en ausencia de la proteína *insulator* es similar a la que se produce en humanos, que no requieren de colesterol en el medio, en la enfermedad neurodegenerativa Niemann Pick tipo C (NPC) (Liscum and Faust, 1987). Se ha descrito que la mutación del gen responsable de esta enfermedad en *Drosophila*, *dnp1*, produce una acumulación de esteroides con el mismo fenotipo que la enfermedad en humanos, impidiendo la correcta síntesis de ecdisona. Los animales no pasan de larva I y no consiguen pupar. El fenotipo es recuperado si disponen en el medio de ecdisona (ya que de esta manera no necesitan sintetizarla), un precursor de la ecdisona derivado del colesterol (sugiriendo que el problema no está en la biosíntesis de la ecdisona, sino en la captación del colesterol) o incluso con un exceso de colesterol (indicando que el proceso de tráfico del mismo puede ser sobrecargado, o que hay otros mecanismos alternativos para

transportarlo) (Huang et al., 2005). Descartamos que CTCF influya en el proceso mediante la regulación transcripcional de *dncp1*, ya que analizamos su nivel de expresión y no está influenciado por CTCF (datos no mostrados).

Por otro lado, también se han descrito fenotipos de acumulación de grasa corporal debidos a la alteración del receptor de ecdisona (Kamoshida et al., 2012). Esta posibilidad tampoco la tenemos en cuenta ya que su expresión no depende de la proteína *insulator* como previamente hemos señalado.

Se ha visto que *nobo*, uno de los genes que están influidos por CTCF, no solo es necesario para la correcta síntesis de ecdisona sino que también se requiere para la normal acumulación de colesterol en la glándula protorácica (Enya et al., 2014), por lo que esta manifestación puede ser la consecuencia de la desregulación de su expresión en ausencia de CTCF, aunque un estudio en mayor profundidad sería necesario.

Hemos demostrado que CTCF participa en la esteroidogénesis de *Drosophila* en dos niveles: por un lado regulando la homeostasis del colesterol en la glándula protorácica y por otro en la activación de las enzimas responsables de la síntesis de la hormona. Si estos dos procesos están relacionados o no necesita un análisis más exhaustivo, pero la necesidad de colesterol junto con ecdisona para una recuperación total de los efectos sugiere que CTCF puede formar parte de un mecanismo molecular que relacione ambos procesos.

## **2. PROTEÍNAS QUE INTERACTÚAN CON CP190**

Cada vez es más evidente que quedan por descubrir proteínas que desempeñan función *insulator*. En vertebrados, clásicamente se ha aceptado que CTCF es la única proteína *insulator* (Wallace and Felsenfeld, 2007). En cambio, recientemente se ha descrito un ortólogo del factor GAGA de *Drosophila* que es capaz de realizar una actividad *enhancer-blocking* independientemente de CTCF (Matharu et al., 2010). Este hecho, junto con el descubrimiento de nuevas secuencias *insulators* de las que aún no se sabe las proteínas con las que interactúan (Gomos-Klein et al., 2007) sugieren que otras proteínas *insulators* pueden existir en vertebrados. *Drosophila* es el animal modelo en el que más proteínas *insulators* se han descubierto, y continúan desvelándose en los últimos años (Aoki et al., 2012, Maksimenko et al., 2015).

Como el objetivo de nuestro grupo de trabajo es la caracterización de los *insulators* de *Drosophila* y las funciones que realizan, en el laboratorio se diseñó un experimento para poder identificar nuevos factores que sean necesarios para realizar esta actividad. Para ello se analizó qué proteínas forman un complejo proteico con CP190. Se escogió CP190, ya que se conocía que está presente en los sitios de unión de otras proteínas *insulators*, tal como CTCF, BEAF, Su(Hw) y GAF, y se le considera una proteína común a todos ellos (Gurudatta and Corces, 2009, Negre et al., 2010, Ahanger et al.,

## **Discusión**

2013). CP190 no tiene capacidad de unirse directamente al DNA, y sin embargo se ha detectado que hay sitios en el genoma en los que hay presencia de CP190 pero no de otra proteína *insulator* hasta ahora conocida. Esto ha llevado a pensar que puede haber otras proteínas que recluten CP190 que se desconocen por el momento (Schwartz et al., 2012).

Nuestro grupo ha purificado este complejo y hemos visto que CP190 interactúa con muchas proteínas (Cuartero et al., 2014). Entre otras, identificamos dos hasta el momento desconocidas (Ibf1 e Ibf2). Estas proteínas localizan en *insulator bodies* y se asocian a la cromatina a lo largo de todo el genoma en sitios donde también hay CP190, pudiendo o no estar unidas a la vez por otras proteínas *insulators*, principalmente CTCF. Forman un heterodímero que media la unión de CP190 en algunos sitios del DNA, por lo que Ibf1 e Ibf2 participan en una subclase de *insulators* de CP190. Además, Ibf1 e Ibf2 son necesarias para la actividad *enhancer-blocking* en el complejo Bithorax, y hemos visto que la mutación de Ibf2 produce transformaciones homeóticas. Por todo ello, hemos descrito un nuevo mecanismo de reclutamiento de CP190 a la cromatina necesario para la actividad *insulator*.

Durante esta tesis he participado en esta investigación en diferentes aspectos. Por un lado, he caracterizado la actividad *enhancer-blocking* de ambas proteínas confirmando su función como *insulators* y, por otro, he detectado la presencia de las dos en los *insulator bodies*, así como la colocalización en los mismos con CP190. He profundizado en el estudio de estos cuerpos para detectar que ambas proteínas se necesitan mutuamente para estar en ellos, pero que la presencia de CP190 es independiente de ambas proteínas. Por último, he comenzado el estudio de otra proteína que habíamos detectado que interactuaba con CP190, pita, aportando los primeros datos que sugieren que forma parte del mismo complejo que las otras tres proteínas, y una posible actividad *enhancer-blocking*. No obstante, esta investigación se detuvo debido a una publicación que detallaba lo que nosotros estábamos estudiando (Maksimenko et al., 2015).

### **2.1. Ibf1 e Ibf2 son proteínas *insulators***

#### **2.1.1. Ibf1 e Ibf2 tienen actividad *enhancer-blocking***

De manera tradicional se ha definido a las proteínas *insulators* por su capacidad para bloquear la comunicación entre elementos reguladores transcripcionales, como *enhancers* o promotores, así como por ejercer un efecto barrera impidiendo la expansión del silenciamiento génico de una zona de heterocromatina a una transcripcionalmente activa, situándose en la frontera entre ambas.

Analizar el funcionamiento de los *insulators* no es sencillo. La función que realizan, separar regiones reguladoras, es un mecanismo de acción dentro del núcleo, y por tanto no está ligado a un grupo de genes en particular o a una vía de señalización que se puedan analizar de manera específica, sino que afecta a regiones muy diversas y los fenotipos no tienen por qué dar indicaciones claras. La manera en que clásicamente se ha caracterizado su función (actualmente está la técnica 3C

y todas las que se han derivado de ella) es en ensayos con constructos transgénicos en los que son capaces de bloquear la comunicación entre un *enhancer* y su promotor y/o evitar la expansión de los efectos del silenciamiento de la cromatina, actuando como barrera (Kellum and Schedl, 1991, Geyer and Corces, 1992, Kellum and Schedl, 1992). No obstante, hay que tener siempre en cuenta que se está investigando en un contexto artificial.

En *Drosophila*, el complejo Bithorax contiene varias secuencias reguladoras de la transcripción que organizan la expresión espacio-temporal de los genes homeóticos que están presentes en la región. Diversos estudios han demostrado que la correcta comunicación entre estas secuencias reguladoras depende de la función de varios *insulators*, como Mcp, Fab6, Fab7 y Fab8 en el caso del gen *Abd-B*, y que CP190 es necesaria para la actividad de estos elementos (Karch et al., 1994, Cavalli and Paro, 1998, Barges et al., 2000, Perez-Lluch et al., 2008).

Para nuestro análisis optamos por el empleo de diversas líneas transgénicas que contienen el *insulator* Fab8 insertado entre el *enhancer* y el promotor del gen *white*. Elegimos realizar así el ensayo por diversos motivos. Por un lado, elegimos el Fab8 porque se había comprobado por ChIP-*seq* que Ibf1 e Ibf2 se unían fuertemente al mismo. Además, diversos estudios constatan la capacidad de este *insulator* para bloquear la comunicación entre *enhancer* y promotor, y ha sido empleado en diferentes investigaciones, tanto *in vitro* como *in vivo* (Barges et al., 2000, Moon et al., 2005, Ciavatta et al., 2007, Bohla et al., 2014).

Todas las líneas transgénicas que utilizamos contienen esta secuencia *insulator* entre el *enhancer* y el promotor del gen *white*. Este gen *reporter* puede que sea el que ha sido más empleado, ya que la variación en la expresión génica da un fenotipo muy obvio y muy sencillo de evaluar. Como no podía ser de otra manera también ha sido empleado para conocer las propiedades *insulators* de diferentes secuencias (Kellum and Schedl, 1992, Hagstrom et al., 1996, Sigrist and Pirrotta, 1997, Gruzdeva et al., 2005).

El lugar de inserción del transgén con el *insulator* influye en estos ensayos ya que analizamos propiedades de la organización de la cromatina, y esta organización puede no ser posible en algunos contextos genómicos. Por ejemplo, se han descrito ensayos de este tipo en que los *insulators* transgénicos pueden bloquear la comunicación a través de la interacción entre ellos, con la formación de un *loop* que supone una barrera física que aísla un elemento regulador del otro (Kyrchanova et al., 2013). La localización en la que se produce la inserción parece influir en el resultado, ya que puede que existan solo unos pocos sitios en el genoma donde los *insulators* endógenos y los transgénicos puedan formar los *loops*. En otros casos, la capacidad para bloquear la comunicación puede depender de la fuerza del *enhancer*, que varía según el sitio de inserción del transgén (Kostyuchenko et al., 2009). Este motivo puede explicar el hecho de que veamos variabilidad en los resultados en función de la línea transgénica analizada. Hay diferencias que son más sutiles y otras más evidentes, problema que se solventa usando varias líneas para evitar de esta manera una conclusión errónea.

## **Discusión**

Por otro lado, realizamos el ensayo con diferentes líneas mutantes de la proteína con el fin de que el efecto percibido no sea debido a mutaciones en el fondo genético. El hecho de que el efecto se constata en diferentes líneas confirma que realmente se debe a la falta de proteína en estudio. Estos análisis se están realizando con la mutación en heterozigosis. Al analizar el resultado en ausencia total de la proteína, vemos que el efecto es incluso mayor, como se aprecia por el incremento de la coloración de los ojos respecto al ensayo realizado con la misma línea mutante en heterozigosis. Las líneas mutantes de *lbf1* e *lbf2* son nulas, por lo que la ausencia total de la proteína incrementa el efecto es acorde a esta carencia de la misma, ya que los individuos heterozigotos aún son capaces de sintetizarla aunque en menor cantidad que los individuos control.

Por último, está el factor de la existencia de una cierta variabilidad entre individuos. Por ello, cuantificamos la pigmentación del color de los ojos con un ensayo previamente ya empleado con este fin (Ephrussi and Herold, 1944, Torras-Llort et al., 2010), evaluando una cantidad suficiente de individuos para que las diferencias se puedan considerar significativas.

Teniendo en cuenta todos estos factores, en todas las líneas analizadas observamos un efecto, por lo que podemos concluir que nuestros análisis muestran que ambas proteínas son necesarias para la actividad *enhancer-blocking* del *insulator* Fab8 del complejo Bithorax.

La desregulación del control de la expresión del complejo Bithorax da lugar a que un segmento presente la identidad de otro, fenómeno conocido como transformación homeótica. En el laboratorio se ha visto que la mutación de *lbf2* produce una transformación similar a la observada en moscas con una delección del Fab8 (Barges et al., 2000, Cuartero et al., 2014), lo que refuerza nuestros resultados obtenidos. No podemos comprobar si esto también sucede con la falta de *lbf1*, ya que los individuos homocigotos no sobreviven hasta fase adulta, pero el hecho de que *lbf1* forme siempre un complejo con *lbf2* sugiere que el papel comprobado para *lbf2* es extrapolable a *lbf1*.

En conjunto, estos resultados nos permiten concluir que *lbf1* e *lbf2* son necesarias para la funcionalidad del elemento Fab8.

### **2.1.2. *lbf1* e *lbf2* se necesitan mutuamente para localizar en los *insulator bodies***

Estudios recientes indican que las funciones de los *insulators* van más allá de sus actividades clásicas como elementos barrera o para bloquear la comunicación *enhancer*-promotor. Se ha propuesto que los *insulators* son capaces de regular la estructura de la cromatina a través de interacciones a larga distancia intra e intercromosomales mediante la formación de *loops*. Estos contactos se pueden hacer con otros elementos reguladores, como *enhancers* y promotores o bien entre ellos. La interacción entre proteínas *insulators* da lugar a la formación de sitios nucleares donde aparecen juntos, denominados *insulator bodies*. Se afirmó que en estos cuerpos se congregan

muchas secuencias *insulators* y el DNA que se localiza junto a ellas. Estas secuencias están en diferentes partes del genoma, y se agrupan mediante interacciones entre las proteínas *insulators*, estabilizando los *loops* que separan dominios de diferente estado transcripcional y, por tanto, controlando la organización y función del genoma (Gerasimova et al., 2000, Labrador and Corces, 2002). Sin embargo, estas estructuras no se entienden correctamente, sobre todo los mecanismos por los que se forman y su función nuclear.

La colocalización de proteínas con actividad *insulator* fue lo que llevó a caracterizar estos cuerpos. En numerosos estudios se han relacionado con sitios en los que Su(Hw), CP190, Mod(mdg4) y CTCF están juntos (Gerasimova and Corces, 1998, Gerasimova et al., 2000, Pai et al., 2004, Gerasimova et al., 2007). Sin embargo, no todas las proteínas *insulators* tienen que colocalizar con todas en estos cuerpos. Recientemente se ha descrito que la proteína BEAF no está en los propios *insulator bodies*, sino que queda formando halos alrededor de ellos (Schoborg et al., 2013). Vemos que las dos nuevas proteínas *insulators* están formando parte de los *insulator bodies*, y que colocalizan entre ellas y con CP190. No obstante, que se aprecien en los mismos cuerpos no implica que formen parte del mismo tipo de *insulator*. CTCF y Mod(mdg4) no colocalizan en cromosomas politénicos, y sin embargo sí que coinciden en *insulator bodies* (Gerasimova et al., 2007). Lo que resulta interesante de la colocalización de Ibf1 e Ibf2 es que se necesitan mutuamente para estar presentes en ellos. La necesidad de una proteína para que otra pueda formar parte de los cuerpos ya ha sido descrita. Por ejemplo, la falta de Mod(mdg4) altera la capacidad de Su(Hw), pero no de CP190 para formar los *insulator bodies*. Tan solo cuando Mod(mdg4) está presente Su(Hw) colocaliza en los cuerpos que tienen CP190, sugiriendo que la interacción entre Mod(mdg4) y Su(Hw), pero no la de CP190 y Su(Hw), es necesaria para que Su(Hw) entre a formar parte de estos agregados (Schoborg et al., 2013). Sin embargo, mutaciones de Su(Hw) no alteran la capacidad ni de CP190 ni de Mod(mdg4) para formar estos cuerpos.

Sería interesante estudiar si Ibf1 e Ibf2 están formando parte de los *insulator bodies* en ausencia de CP190, CTCF y/o Su(Hw). Nuestro grupo ha comprobado que la mayoría de los sitios en los que CP190 requiere del dímero para unirse al DNA son aquellos en los cuales colocalizan con CTCF, mientras que en las regiones en las que la unión de CP190 no depende de las Ibfs se corresponden con sitios donde interacciona con Su(Hw). Si al eliminar estas proteínas siguen estando presentes en los *insulator bodies* sugeriría que son independientes de ellas, y constituirían un nuevo tipo de *insulator body*. En el caso de no formar parte de ellos en ausencia de CTCF o Su(Hw) indicaría que la presencia de Ibf1 e Ibf2 en los cuerpos se realiza a través de interacciones con una de estas proteínas, aunque según nuestros resultados previos es poco probable que dependa de Su(Hw). Los estudios iniciales habían establecido que CP190 era necesaria para que las proteínas *insulators* estuvieran formando parte de los *insulator bodies*, y que su ausencia implicaba la disgregación proteica (Gerasimova et al., 2000). Sin embargo, estudios recientes señalan que la falta de CP190 no afecta a la capacidad de Mod(mdg4) ni de BEAF-32 para formar cuerpos (Schoborg et al., 2013), hecho sorprendente por la colocalización de Mod(mdg4) y CP190 (Pai et al., 2004) y BEAF-32 y CP190

## **Discusión**

en múltiples sitios del genoma (Bushey et al., 2009), por lo que resulta interesante estudiar el comportamiento de Ibf1 e Ibf2 en ausencia de CP190.

A pesar de los avances en el conocimiento de los *insulators*, todavía no es evidente la función que tienen los cuerpos nucleares que forman. Estudios iniciales aseguraban que tenían un papel funcional en la organización de la cromatina y que su formación guardaba relación con la regulación transcripcional (Gerasimova et al., 2000, Byrd and Corces, 2003). Sin embargo, unos años más tarde se propuso que los *insulator bodies* eran independientes de las secuencias *insulators*, y por tanto, de su actividad en la organización de la cromatina, considerándose como meros depósitos de proteínas (Golovnin et al., 2008). Por otro lado, un estudio reciente afirma que son cuerpos nucleares que se forman por estrés osmótico y muerte celular (Schoborg et al., 2013), que no están unidos a la cromatina y, por tanto, carentes de actividad *insulator*.

Si estas formaciones no están desempeñando la actividad que siempre se les había asignado, ¿cuál es la función de estas agregaciones? Lo cierto es que la organización proteica por estrés osmótico afecta de manera específica a las proteínas *insulators*, sin alterar a otras proteínas de unión al DNA, como las que forman parte de los *Polycomb bodies*. El significado fisiológico de la separación de la cromatina y la formación de estos cuerpos no se conoce. El hecho de que otras proteínas con unión al DNA no sufran esta reestructuración y dado el papel central de los *insulators* en la organización de la cromatina, los *insulators* deben estar realizando alguna función en estos cuerpos.

Quizás la separación de las proteínas *insulators* del DNA se necesite para alterar o reorganizar los dominios de cromatina como respuesta al estrés, ya que no se ha confirmado una eliminación total de las proteínas *insulators* de la cromatina, sino una disminución en la unión en los sitios analizados (Schoborg et al., 2013). Las que quedan unidas al DNA continuarían mediando su acción. Recientemente un estudio ha analizado los cambios que se producen en la distribución de las proteínas *insulators* bajo estrés térmico (Li et al., 2015). Esta alteración origina un cambio en su localización, lo cual supone una reorganización de la estructura de la cromatina. Aunque se afirma que los *insulator bodies* no se forman bajo estrés térmico, puede que las proteínas que se mantienen unidas se reorganicen bajo los dos tipos de estrés, contribuyendo a la regulación de la organización de la cromatina, y que los *insulator bodies* realmente sean solo depósitos proteicos.

Si la teoría de que se forman bajo osmoestrés es cierta, no debiéramos observar estos *insulator bodies*, pero lo cierto es que los vemos en las diferentes preparaciones analizadas. Una posibilidad podría ser que la técnica a la que la muestra está sometida genere este tipo de estrés. Por ejemplo, el tipo de tampón de disección empleado así como el tiempo que se mantiene en él hasta la fijación podría suponer un estrés para el tejido y en respuesta generara *insulator bodies*.

Todavía queda mucho por investigar sobre cómo y porqué se forman los *insulator bodies*, si pueden contribuir en las funciones de los *insulators* y las consecuencias de su formación en la organización nuclear y dinámica del genoma. Sería interesante realizar *ChIP-seqs* para analizar los

cambios globales en la unión al DNA de las proteínas *insulators* y técnicas 3C para evaluar los contactos mediante *loops* que conlleva la formación de estos cuerpos, así como valorar la repercusión que tienen a nivel transcripcional.

Las observaciones de que Ibf1 e Ibf2 son capaces de bloquear la comunicación entre *enhancer* y promotor, que colocalizan con otras proteínas *insulators* en los *insulator bodies* del núcleo, así como por otros resultados obtenidos fuera de esta tesis como su capacidad de unirse al DNA mediando la asociación de CP190 en numerosos sitios a lo largo de todo el genoma, consideramos que ambas son proteínas *insulators*.

No obstante, se ha de estudiar en más profundidad de qué manera estas proteínas contribuyen al funcionamiento de los *insulators*. Para comenzar, se ha de investigar cómo se produce la interacción con CP190. Se cree que CP190 hace de engranaje entre las proteínas *insulators* unidas al DNA y que esto genera *loops* y barreras entre regiones reguladoras. Si éste fuera el caso, podría ser que Ibf1 e Ibf2 se unieran al DNA a través del reconocimiento de su motivo consenso y que mediante la interacción con CP190 acabaran formando *loops* entre diferentes lugares que tienen estos motivos. Otra posibilidad podría ser que los contactos también se dieran con CTCF, y que las Ibf interaccionaran con CTCF y CP190. Incluso puede que CP190 se una a otros sitios a través de otras proteínas, y que se forme un *loop* con los sitios de unión de Ibf1 e Ibf2 a través de la interacción entre las tres. Independientemente de cómo se establezca este contacto entre ellas sería muy informativo hacer experimentos de ChIA-PET, en el que se determina la contribución de una proteína en concreto en la formación de redes de contacto entre secuencias. Otra posibilidad sería el Hi-C para comparar las interacciones intra e intercromosómicas entre una situación normal y la mutación de cada una de las proteínas *insulators*, para, de esta manera, determinar cuál es la contribución de cada una a la formación de contactos entre secuencias.

## **2.2. Pita como *insulator***

Las primeras evidencias que hemos obtenido al realizar el ensayo para evaluar la capacidad *enhancer-blocking* de la proteína pita sugieren que interviene en la actividad *insulator* del Fab6. No obstante, estos resultados incipientes no son concluyentes, pues como hemos señalado previamente es necesario un ensayo exhaustivo y mucho más amplio para poder sacar conclusiones. Además, la única comprobación que hemos hecho de su función sobre Fab7 indica que no está realizando ninguna función en este *insulator*. No obstante, si para afirmar que ejerce una acción son necesarias varias comprobaciones, lo mismo se aplica para rechazar su función. Según el artículo publicado por el grupo de Georgiev (Maksimenko et al., 2015) pita es una proteína *insulator* con actividad *enhancer-blocking*, al menos sobre el *insulator* Mcp, donde colabora con CTCF y CP190. Esto, junto con nuestros resultados que indican que participa en la actividad del *insulator* Fab6, sugiere que pita



## **Discusión**

participaría en la actividad *enhancer-blocking* de los diferentes *insulators* de la región reguladora de *Abd-B*, de manera similar a como lo hacen otras proteínas *insulators* como CP190, CTCF, Ibf1 e Ibf2.

Hemos comprobado que Ibf1 e Ibf2 interaccionan juntas con CP190, y por otro lado, que Pita lo hace con CP190. Podría darse el caso de que las detectásemos en el mismo complejo multiproteico no por una relación directa entre las Ibf y Pita, sino porque CP190 fuera capaz de interaccionar con los dos grupos de proteínas, como en el caso de CTCF y Mod(mdg4), que se encuentran en los mismos *insulators bodies* gracias a la interacción de ambas por separado con CP190, pero que no lo hacen entre ellas (Gerasimova et al., 2007). Sin embargo, rechazamos esta posibilidad ya que tanto Ibf1 como Ibf2 coimmunoprecipitan con Pita. Para comprobar si realmente se unen a los mismos sitios del genoma se podría realizar una inmunotinción en cromosomas politénicos de las proteínas y comparar su lugar de unión. O mucho más informativo incluso, un ChIP-seq para poder comparar todos los sitios a los que se unen a lo largo del genoma.

Con los datos iniciales que se obtuvieron de Pita se pueden señalar algunas similitudes entre ésta y otras proteínas *insulators* además de su actividad *enhancer-blocking*. Ya en el año 2003 (Laundrie et al., 2003) se señaló que Pita pudiera regular la transcripción génica, y fue confirmado dos años después. Se vio que en ausencia de la proteína, 33 genes se reprimían mientras que otros 10 aumentaban su expresión (Page et al., 2005). Esta observación es consistente con el efecto sobre la regulación transcripcional que ejercen otras proteínas *insulators* como CTCF, CP190, Su(Hw), Ibf1 e Ibf2, que regulan la expresión de genes, activando unos y silenciando otros (Bartkuhn et al., 2009, Schwartz et al., 2012, Cuartero et al., 2014).

Por otro lado, la ausencia de Pita en línea germinal produce defectos en el desarrollo de las cámaras ováricas, con núcleos de las células nodrizas anormales o ausentes, que conlleva la muerte celular prematura (Laundrie et al., 2003). Este efecto es consecuencia del fallo en la endoreduplicación de las células nodrizas. Pita no realiza esta función directamente, sino mediante la regulación de la expresión del gen *Orc4*, implicado en la replicación del material genético (Page et al., 2005). Se ha visto que la ausencia de la proteína Su(Hw) también causa defectos en la ovogénesis así como esterilidad en las hembras (Parkhurst et al., 1988). Como en el caso de Pita, Su(Hw) media esta acción a través del silenciamiento transcripcional que regula (Schwartz et al., 2012). Sin embargo, ésta no es una función compartida con otras proteínas *insulators* como CP190 o Mod(mdg4) (Baxley et al., 2011), y parece independiente de su actividad *enhancer-blocking*. Previamente también hemos señalado que CTCF no necesita a CP190, Ibf1 ni Ibf2 para mediar su acción como regulador transcripcional. Por tanto, las proteínas *insulators* pueden controlar la expresión génica independientemente de CP190. Sería interesante estudiar si Pita requiere de CP190 para mediar esta función o bien ocurre lo mismo que con Su(Hw) y CTCF.

La muerte celular antes de la ovogénesis tardía, como sucede en ausencia de Pita, está regulada por ecdisona (Buszczak et al., 1999, Carney and Bender, 2000). Los mutantes de genes de respuesta a ecdisona muestran degeneración prematura de la cámara ovárica similar a la que produce la

ausencia de pita. De hecho, la letalidad en larva que genera la falta de ésta puede ser recuperada si se añade ecdisona al medio (Laundrie et al., 2003), sugiriendo que pita pudiera regular o interactuar con componentes de la vía de síntesis de la ecdisona. Los resultados que hemos obtenido de que CTCF está regulando también esta vía indican que no sería excepcional que una proteína *insulator* influyera en el proceso. No obstante, queda por determinar qué función desempeña pita en relación con la síntesis de ecdisona. Ya que muchos de los sitios de unión de pita a lo largo del genoma están también enriquecidos en CTCF (Maksimenko et al., 2015), podría resultar de interés evaluar si ambas proteínas están relacionadas a la hora de mediar esta acción.

### **3. NUEVAS PROTEÍNAS CON ACTIVIDAD *INSULATOR*: Haspin**

En nuestra búsqueda de nuevas proteínas *insulators* hemos realizado un *screening* para detectar proteínas que ejerzan función *enhancer-blocking*. Hemos obtenido varios candidatos con posible actividad *insulator*, y nos decantamos por el estudio de Haspin. Esta proteína está altamente conservada, pues no solo está presente en nuestro animal modelo sino que también se ha visto en otros como *C. elegans*, *A. thaliana*, *S. pombe*, *S. cerevisiae*, e incluso en vertebrados como ratones (donde se descubrió (Tanaka et al., 1999)) o en humanos (Higgins, 2001). Su dominio C-terminal, el que tiene actividad kinasa, presenta gran homología en todas las especies, aunque no tanto el que le permite el reconocimiento de los sustratos, el N-terminal. Haspin ha sido muy estudiada por su actividad durante mitosis para la localización de Aurora B en los centrómeros (Yamagishi et al., 2010, Trivedi and Stukenberg, 2015). Sin embargo si juega un papel durante interfase no se conoce. Existen otros ejemplos de proteínas con función en ambas fases y entre ellas proteínas que participan en la organización de la cromatina y que han sido caracterizadas como proteínas *insulators*. Uno de ellos es el factor de transcripción TFIIC, que se une a centrómeros durante mitosis (Iwasaki et al., 2010), y recientemente se ha clasificado como proteína *insulator* (Van Bortle and Corces, 2012, Kirkland et al., 2013). Éste no es un caso aislado; otro ejemplo es el complejo de las cohesinas, cuya participación en la segregación del material genético así como en la organización de la cromatina y en la regulación de la expresión génica ya ha sido explicado en la introducción (apartados 1.1.1.3., 1.2.2.1., 1.2.2.3. y 3.1.1.).

En esta tesis hemos descrito el papel de Haspin como proteína con actividad *insulator* por su capacidad para bloquear la comunicación entre un *enhancer* y un promotor, pero especialmente por su papel en la organización y estructura de la cromatina, proponiendo que realiza sus acciones a través de su dominio kinasa. Además, hemos analizado su relación con proteínas del grupo Polycomb, el complejo de las cohesinas así como una posible relación antagónica con el de las condensinas.

### **3.1. Haspin tiene actividad *enhancer-blocking***

Teniendo en cuenta todas las premisas que hay que valorar a la hora de realizar un ensayo de *enhancer-blocking*, expuesto en el apartado 2.1.1. de esta discusión, podemos concluir que Haspin es necesaria para la correcta función de los *insulators* de la región reguladora del gen *Abdominal-B*. El *screening* sugería que Haspin estaba involucrada en la actividad *insulator* del elemento Fab7. Esto lo hemos demostrado al utilizar diferentes líneas transgénicas y diferentes mutantes de la proteína. Además, no solo es importante para Fab7, sino que también para otros *insulators* de la región reguladora del gen *Abd-B* como Fab6 y Fab8.

El evidente papel que juega Haspin en la regulación de esta región se contradice con el hecho de no observar transformaciones homeóticas en ausencia de la proteína. Este fenotipo se ha visto en la falta de cualquiera de los tres Fabs que hemos analizado (Barges et al., 2000, Zhang et al., 2003, Iampietro et al., 2010), así como en ausencia de algunas de las proteínas que regulan su expresión, como CTCF o Ibf2 (Mohan et al., 2007, Cuartero et al., 2014). Una consecuencia que habría de tener la falta de Haspin si es requerida para la correcta expresión de esta región habría de ser, por tanto, la de presentar este tipo de transformaciones. No obstante, la ausencia de la proteína no genera este efecto. Es posible que la manera en que Haspin afecta a la funcionalidad del elemento sea diferente a la de las otras, no produciéndose la misma manifestación.

La falta de otras proteínas *insulators* produce letalidad, como en el caso de CP190 (Butcher et al., 2004) o CTCF (Mohan et al., 2007). Nosotros lo que observamos en mutantes de *Haspin* es que el tiempo de supervivencia de los individuos es significativamente menor en ausencia de la proteína, sin tener afectación alguna sobre la viabilidad. Esto es acorde a lo publicado en la literatura, donde la falta de Haspin tampoco provoca la muerte en otros organismos como *S. pombe* (Yamagishi et al., 2010).

### **3.2. Haspin es necesaria para la organización de la cromatina en el núcleo en interfase**

La visión tradicional de los *insulators* como bloqueadores de la comunicación entre *enhancer* y promotor, y como barrera para la expansión del silenciamiento por parte de la heterocromatina es demasiado reducida para explicar diversas observaciones recientes. Se ha demostrado que los *insulators* también pueden favorecer la comunicación *enhancer*-promotor directamente para estimular la expresión génica (Erokhin et al., 2011, Xu et al., 2011, Eldholm et al., 2014). Esta apreciación parece ser antagónica a sus funciones clásicas, por lo que debe existir un mecanismo que dé una explicación que permita entender todas sus actividades. En una visión global, los *insulators* interaccionan entre ellos formando *loops* que organizan la estructura de la cromatina. De hecho,

según la localización relativa de los *insulators* entre ellos en esta organización en 3 dimensiones o respecto a otros elementos de la regulación génica con los que también pueden contactar, estos *loops* pueden favorecer o bloquear la comunicación entre *enhancer* y promotor, ejerciendo como barreras físicas para la misma. De esta manera, su función como reguladores de la estructura de la cromatina explica todas las actividades que desempeñan a nivel nuclear.

No hemos podido detectar Haspin con el anticuerpo que generamos ni en interfase ni en mitosis a pesar de sobreexpresar la proteína, por lo que hemos caracterizado Haspin a través de la proteína tagueada. Recientemente ha sido publicado un estudio en el que observan la proteína endógena en ratones durante meiosis (Wang et al., 2015), pero la mayoría de estudios sobre Haspin están basados en resultados obtenidos al caracterizar la proteína tagueada en lugar de la endógena (Dai et al., 2005, Yamagishi et al., 2010, Nguyen et al., 2014). Se ha descrito que tiene presencia nuclear a lo largo de todo el ciclo celular (Dai et al., 2005) y mediante tinción de núcleos interfásicos confirmamos que durante interfase también se encuentra aquí en *Drosophila*, concretamente en la periferia nuclear. Es interesante que incluso antes de que se conociera que Haspin era la kinasa que fosforila H3T3 se había dicho que la proteína que lo hiciera estaba asociada a estructuras periféricas ya que la fosforilación de este residuo en profase se iniciaba en las proximidades de la envoltura nuclear y además se detectaba en preparaciones que contenían esta región (Polioudaki et al., 2004). A pesar de su presencia nuclear no la vemos unida al DNA como se ha publicado (Dai et al., 2005). Puede que la causa de esto sea que Haspin establece el contacto con la cromatina para ejercer su actividad con una dinámica rápida y/o con muy poca frecuencia por lo que mediante microscopía no seamos capaces de detectar esta interacción. Es posible que este problema se pueda solventar mediante la mutación del dominio kinasa. Planteamos esta posibilidad ya que ha sido publicado que, en ratones, Haspin se une al DNA con mayor afinidad cuando no es capaz de autofosforilarse (Tanaka et al., 1999). Por tanto al eliminar o mutar su dominio kinasa en *Drosophila* puede que también tenga este efecto.

Una de las formas de organización de la cromatina es mediante la asociación con diversas estructuras nucleares. Haspin, a través de la interacción con zonas periféricas, participaría en la compartimentalización nuclear. Esto no solo lo afirmamos en base a los resultados obtenidos en la inmunolocalización de la proteína, sino que además hemos visto la importancia de esta proteína para mantener una correcta estructura nuclear, pues su ausencia implica un drástico cambio fenotípico con fuertes invaginación en la periferia nuclear. El hecho de que no afecte a la cantidad de lamina que se expresa, nos permite concluir que está realizando un papel estructural. Las zonas de la periferia nuclear que contactan con las proteínas de la lamina unen cromatina con poca cantidad de genes y transcripcionalmente silenciados (van Bemmelen et al., 2010). Sin embargo, otras estructuras, los complejos de los poros nucleares, unen genes transcripcionalmente activos, como se ha descrito para ciertos genes, como los del cromosoma X en machos de *Drosophila* (Akhtar and Gasser, 2007).

Las proteínas *insulators* son las responsables de la separación de las regiones periféricas con diferente estado transcripcional. CP190 está unida a la cromatina, pero hemos visto que su presencia en la periferia nuclear aumenta en ausencia de Haspin. Nuestros resultados apuntan a que es un

## **Discusión**

efecto exclusivo sobre CP190, ya que no repercute en la localización de otras proteínas *insulators* como Ibf2 o BEAF, aunque no podemos rechazar que tenga efecto sobre otras. Esta duda podría quedar solventada mediante la realización de un ChIP-seq analizando los cambios en la localización de las proteínas *insulators* en ausencia de Haspin, si es que se dan. ¿Qué es lo que provoca esta alteración en la localización de CP190? No tenemos ningún indicio de qué puede estar sucediendo. Es posible que Haspin regule de alguna manera la función de las proteínas *insulators* en esta región, o simplemente que este cambio que observamos en CP190 sea una consecuencia de la remodelación de la estructura nuclear que se produce en ausencia de Haspin.

### **3.3. Haspin ejerce su acción a través de su dominio kinasa**

El dominio C-terminal de Haspin, el que tiene la actividad kinasa, es el que está conservado a lo largo de la evolución. Es funcional durante todo el ciclo celular, ya que presenta una conformación intrínsecamente activa (no tiene ninguna zona para su activación mediada por fosforilación, un mecanismo muy frecuente en el control de la actividad kinasa) (Eswaran et al., 2009, Villa et al., 2009). De hecho, desde el comienzo de su estudio se propuso que su dominio kinasa estaba siempre activo pues la sobreexpresión de la proteína conllevaba la fosforilación de H3T3, único sustrato que se conoce de Haspin hasta el momento, durante todo el ciclo celular (Dai et al., 2005). Sin embargo, recientemente se ha señalado que, junto a este dominio kinasa, tiene un dominio básico autoinhibitorio, que bloquea el dominio catalítico de manera alostérica durante interfase, impidiendo su acción. Siendo, por tanto, solo activo durante la fase divisoria (Ghenoiu et al., 2013). Nosotros hemos confirmado que, al menos en *Drosophila*, este dominio es activo en interfase, como se observa por la fosforilación de H3T3ph en cromosomas politénicos mediada por Haspin. Además, los efectos que la proteína presenta sobre la condensación de la cromatina dependen de que su dominio kinasa sea funcional. Nuestros resultados son más acordes con la propuesta de un mecanismo en el que el dominio kinasa pase de manera dinámica de un estado unido o libre del dominio autoinhibitorio, lo que permite la fosforilación por parte de la kinasa, aunque con una cinética más lenta a como lo hace en mitosis.

Está descrito que Haspin fosforila H3T3 en mitosis, y que la marca es esencial para las labores que desempeña en esta fase del ciclo celular, como la localización de Aurora B en los *inner centrómeros* (Yamagishi et al., 2010) o la división asimétrica en células pluripotenciales (Tran et al., 2012). Mediante inmunotinción de células mitóticas hemos confirmado que Haspin también fosforila H3T3 durante mitosis en *Drosophila*. Nuestros resultados muestran la total desaparición de la señal en centrómeros en ausencia de Haspin, lo que indica que Haspin si no es la única, al menos es la principal kinasa que media esta marca. Si existe otra kinasa que sea capaz de producir la modificación en muy pequeña cantidad no lo podemos advertir de esta manera. En mamíferos, VRK-1 fosforila la histona en este residuo junto con Haspin (Kang et al., 2007, Valbuena et al., 2011, Jeong et al., 2013).

Puede que NHK-1, su ortólogo en *Drosophila*, también produzca esta PTM por la homología en el dominio kinasa y en el del reconocimiento de sustratos de ambas proteínas.

Por otro lado, también hemos visto que realiza la fosforilación de H3T3 en interfase. Este fenómeno es muy interesante ya que no se tenía constancia previamente de la presencia de esta marca en la fase no divisoria. Inicialmente se asoció la fosforilación de las histonas como una marca exclusivamente mitótica, pero con el paso del tiempo se ha visto que no solo provoca la condensación de los cromosomas sino que también media en procesos interfásicos. Muchos son los estudios que asocian marcas como H3S10ph y H3T3ph con la condensación de los cromosomas mitóticos, pero cada vez hay más evidencia del papel que juega H3S10ph en interfase, y es que la fosforilación de este residuo por parte de Jil-1 es necesaria para la expresión de los genes del cromosoma X en machos de *Drosophila* (Wang et al., 2001, Deng et al., 2008, Regnard et al., 2011).

Ya que no podemos detectar Haspin unida a la cromatina, optamos por estudiar su localización en interfase a través de H3T3ph, pues, como hemos señalado, es la proteína principal que se encarga de realizar esta marca posttranscripcional. Mediante inmunolocalización en cromosomas politénicos vemos que H3T3ph se sitúa en interbandas, es decir, zonas transcripcionalmente activas. De hecho, se puede concretar que incluso la señal que observamos es mayor en los *puffs*, zonas muy activas transcripcionalmente. Marcas epigenéticas relacionadas con la activación de la expresión génica, como H3S10ph, también se localizan en interbandas (Jin et al., 1999). De hecho, el que pueda favorecer la transcripción en glándulas salivares es consistente con nuestro resultado sobre el papel de Haspin en la condensación. Hemos visto que Haspin es necesaria para la descondensación de la cromatina interfásica, manteniendo un estado abierto del material genético, lo que facilita la accesibilidad de la maquinaria transcripcional a los genes. No obstante, no tenemos indicaciones de que influya en la transcripción génica en glándulas salivares. No hemos podido comprobar que en cromosomas politénicos H3T3ph colocalice con la RNA polimerasa II, ya que los anticuerpos de los que disponemos para detectar ambas proteínas son del mismo animal.

Mediante la misma técnica se ha detectado que muchas de las proteínas *insulators* como CP190, CTCF, Su(Hw) y Mod(mdg4) (Whitfield et al., 1995, Gerasimova et al., 2007) se encuentran en las fronteras entre bandas e interbandas, evitando de esta manera la propagación del silenciamiento de la heterocromatina a la transcripcionalmente activa. La cantidad de bandas que se aprecian en nuestras tinciones de cromosomas politénicos no son muchas, y puede que, debido a la baja resolución de la técnica, no estemos localizando bien esta señal, no distinguiendo si están en interbandas como otros factores transcripcionales o delimitando ambos tipos de cromatina. Ya que CP190 es un cofactor común de todos estos complejos y colocaliza parcialmente con cada una de las proteínas *insulators* (Gerasimova et al., 2007) podría aclararnos este asunto realizar la cotinción de ambas, pero no hemos podido ya que los anticuerpos contra el cofactor y la PTM de los que disponemos son de conejo.

Las glándulas salivares son un tipo celular peculiar, en el que el DNA sufre sucesivas fases de replicación sin realizar mitosis entre ellas. Tratamos de caracterizar H3T3ph en células diploides

mediante ChIP-*seq*. Esta técnica está siendo en los últimos años muy empleada en el campo de estudio de la cromatina, debido sobre todo a los grandes avances que han experimentado las tecnologías de secuenciación. A lo largo del genoma se encuentra principalmente en zonas de heterocromatina pericentromérica, resultado acorde a su función durante mitosis (Yamagishi et al., 2010). Para realizar la técnica hemos partido de un cultivo celular en fase exponencial, es decir, con células en todas las fases del ciclo celular, algunas de ellas en mitosis. Para estudiar la localización de H3T3ph en interfase hubiera sido mejor que estas células estuvieran en fase estacionaria, para reducir el número de divisiones y poder evaluar de manera más exacta su posición en interfase, ya que ahora no podemos decir si la localización en heterocromatina es solo debida a las células en mitosis o a una suma de mitosis e interfase. De todas formas, detectamos que no solo se une aquí, sino que también está enriquecida en otras zonas, que representan grandes áreas del genoma. Al analizar las regiones en las que H3T3ph está presente podemos sacar varias conclusiones; en muchos casos está en zonas delimitadas por proteínas *insulators*, como hemos visto para CP190 e Ibf2. Hemos de señalar que su patrón de unión al DNA difiere bastante del de otras proteínas *insulators* ya que muchas de ellas se localizan en regiones muy concretas del genoma. La localización de H3T3ph flanqueada por estas proteínas pudiera ser una forma de colaborar con ellas en la delimitación de zonas con distinto estado de expresión. Esto se ha aplicado, por ejemplo, a cohesinas, que parece que demarcan zonas transcripcionalmente activas, colaborando con las fronteras que delimitan su sitio de unión para evitar de esta manera la expansión del silenciamiento génico (Dorsett, 2009).

El análisis de los datos aportados en el ChIP-*seq* en función del estado de la cromatina confirma que las regiones de H3T3ph coinciden con zonas de heterocromatina pericentromérica como ya hemos señalado, pero también con otras zonas que están transcripcionalmente silenciadas. Analizamos en detalle la unión de otras proteínas en estas zonas enriquecidas en H3T3ph y comprobamos que no están unidas por proteínas que indican transcripción activa, tal como es el caso de la RNA polimerasa II, y en cambio contienen marcas silenciadoras como H3K27me3 o coinciden con zonas asociadas a lamina.

¿Cuál es la función de H3T3ph dentro de estas zonas silenciadas? Ya en experimentos iniciales sobre fronteras entre los grandes dominios que dividen el genoma se apuntó que estaban enriquecidas en proteínas *insulators*. De hecho la fuerza de estos bordes es directamente proporcional a la cantidad de estas proteínas que se encuentran en ella (Li et al., 2015). Estos grandes dominios son los que determinan la organización del genoma. Se vio que al alterar los bordes, los TADs se fusionaban solo parcialmente, sugiriendo que dentro de los mismos debieran existir factores involucrados en mantener esta separación cuando no existe tal frontera (Nora et al., 2012). Se ha confirmado que dentro de los grandes dominios se producen contactos entre *insulators* a escala más pequeña, que son los que determinan la actividad génica. Sería interesante conocer si estas amplias zonas donde detectamos H3T3ph se corresponden con dominios silenciados y si es necesaria la modificación de la histona para establecer y/o mantenerlos.

El hecho de que la ausencia de Haspin se acompañe de una disminución en la unión de Polycomb en regiones silenciadas parece indicar que la presencia de Haspin es necesaria para la localización de

los complejos Polycomb y mediar el silenciamiento, lo cual explica la colocalización de H3T3ph y H3K27me3. De hecho, no sería de extrañar la relación entre Haspin como proteína *insulator* y su contribución a la represión mediada por las proteínas del grupo de Polycomb. Las interacciones a larga distancia que se realizan entre zonas transcripcionalmente silenciadas no las establecen los PREs, sino las proteínas *insulators* (Li et al., 2011). Sería interesante estudiar si Haspin media en estos contactos entre zonas silenciadas. Con nuestros datos no lo podemos concretar, pero podría ser estudiado con la técnica del ChIA-PET, que permite el análisis de las interacciones a larga distancia que realiza una proteína en concreto. Otra posibilidad es realizar un Hi-C comparando las interacciones entre diferentes regiones en ausencia y presencia de Haspin, pudiendo detallar cuál es su función en la organización de la cromatina.

Se ha demostrado que H3T3ph y H3K4me3, marca que indica activación transcripcional, están relacionadas negativamente, lo cual está de acuerdo con el hecho de que no veamos H3T3ph en zonas activas. Por un lado, se ha afirmado que la presencia de H3K4me3 modula la fosforilación de la H3T3, ya que dificulta que el dominio sea reconocido por Haspin (Eswaran et al., 2009) y se ha visto que la presencia de H3T3ph impide que proteínas necesarias para la transcripción se unan a H3K4me3 (Markaki et al., 2009, Garske et al., 2010). No obstante, ambas marcas existen a la vez *in vivo* (Markaki et al., 2009). Este hecho no ha de sorprender ya que se ha visto la colocalización de otras marcas que se habían considerado que no se pueden producir en *cis*, tal es el caso de H3K9me3 y H3S10ph durante mitosis (Rea et al., 2000) o la coincidencia en promotores de H3K4me3 y H3K27me3 (Voigt et al., 2013, Denissov et al., 2014). Puede que Haspin esté relacionado tanto con la activación como con la represión transcripcional. Las proteínas *insulators* no están funcionalmente ligadas con ninguna de las dos acciones. Por tanto, tanto pueden favorecer la expresión como el silenciamiento génico, dependiendo del gen sobre el que actúen. No hemos analizado la posible relación entre H3T3ph y H3K4me3, por lo que sería interesante comprobar la variación de los niveles de la marca de trithorax así como de H3K27me3 al eliminar la proteína Haspin.

### **3.4. Mecanismo de acción de Haspin**

Nos planteamos si Haspin puede mediar su acción en la organización nuclear de manera análoga o cooperando con Jil-1 por las similitudes apreciadas entre ambas proteínas y por la relación que H3T3ph tiene con H3S10ph. Mientras que esta marca la realiza Aurora-B en mitosis (cuya localización depende de Haspin como hemos señalado en el apartado 3.1 de la introducción), en interfase es producida por Jil-1. Se ha visto que Jil-1 tiene efecto *insulator* impidiendo la propagación de la cromatina silenciada (Ebert et al., 2004) y tanto la proteína como H3S10ph, a través de la cual ejerce sus acciones, están presentes en interbandas en cromosomas politénicos, lugar en el que también detectamos H3T3ph. Jil-1 mantiene un estado abierto de la cromatina descondensando el DNA interfásico (Deng et al., 2008) y el efecto que genera su falta sobre la condensación de los núcleos en interfase es la misma que produce la ausencia de Haspin. Ya hemos señalado que Jil-1 es esencial para la activación de los genes del cromosoma X (Regnard et al., 2011) y que estos se encuentran en



## **Discusión**

la periferia nuclear cuando están activos. Sería interesante conocer si Haspin, que también se encuentra en esta localización en interfase, regula su expresión. No obstante, la falta de Jil-1 provoca una disminución general de la expresión génica. No parece que Haspin afecte de manera general a la expresión génica, pero sería necesario un array para confirmarlo.

Por otro lado, remodeladores de la cromatina, como dMi2 (Fasulo et al., 2012), también tienen la aparente dualidad de acciones que presenta Haspin. De hecho, ambas proteínas se asemejan en sus funciones como organizadores de la estructura del material genético. Se ha visto que dMi2 participa en la represión transcripcional de los genes homeóticos (Kehle et al., 1998) pero que también regula la condensación de la cromatina interfásica sin afectar a la replicación del DNA y localiza en zonas transcripcionalmente activas en cromosomas politénicos, indicando que su papel no está limitado a una de las dos acciones. A pesar de todas estas similitudes, se ha visto que dMi2 también es requerida para la correcta expresión de los genes *heat-shock* en *Drosophila* (Murawska et al., 2011), mientras que Haspin no parece repercutir en la transcripción génica tras un choque de temperatura (datos no mostrados).

El análisis de datos aportados por Hi-C han desvelado que las cohesinas median interacciones a larga distancia entre sus sitios de unión formando *loops*, los cuales involucran elementos regulatorios como *enhancers*, promotores o *insulators*, favoreciendo o inhibiendo la expresión génica (Seitan and Merckenschlager, 2012). No solo las proteínas del propio complejo participan, sino que también intervienen las que están relacionadas con su dinámica de unión (Rollins et al., 1999).

No estaba muy claro cómo afecta pds5 a la estructura de la cromatina, ya que se observó que el mutante nulo de la proteína producía condensación mientras que la carencia de solo una parte de la proteína daba lugar a unos cromosomas politénicos más delgados (Dorsett et al., 2005). El efecto que nosotros observamos es acorde a lo descrito para el mutante nulo de pds5, ya que hemos comprobado que la falta de pds5 provoca condensación de los núcleos en interfase, mismo fenotipo que provoca la ausencia de Haspin. El hecho de que observemos el mismo efecto en ausencia de Haspin y de pds5 nos hace pensar que, como en mitosis, una necesite a la otra. Comprobamos que efectivamente existe una relación entre pds5 y Haspin para mediar la organización del material genético, ya que ésta no es capaz de producir la descondensación de la cromatina en ausencia de pds5. Este resultado es consecuente con los datos de mitosis que reflejan que Haspin necesita de pds5 para su localización en la cromatina. No podemos comprobar si realmente la unión de la kinasa depende de pds5 ya que no la detectamos mediante inmunolocalización. Está descrito que la localización de pds5 en el DNA es independiente de Haspin (Yamagishi et al., 2010). Nuestros resultados afirman que la unión de pds5 al DNA necesita de Haspin, ya que la ausencia de ésta impide que pds5 se una a los cromosomas politénicos. Rechazamos que este efecto sea producido por la alteración de la estructura de la cromatina que induce la ausencia de Haspin ya que no afecta a la unión de otras proteínas. El hecho de que Haspin no influya en la transcripción de su gen, pero disminuya notablemente la cantidad de proteína indica que la ausencia de Haspin induciría la degradación de pds5 al no poderse unir al DNA, como ocurre con otras proteínas. Se ha visto que,

por ejemplo, proteínas de PRC2 se degradan cuando no se puede formar el complejo (Tan et al., 2007). De esta manera podemos establecer la hipótesis de que pds5 y Haspin se necesitan mutuamente para sus efectos en la estructura de la cromatina.

No hay evidencias directas pero se ha sugerido que Haspin no afecta a la unión de las cohesinas, sino más bien a su mantenimiento una vez unidas al DNA en mitosis. Acorde a esto hemos detectado que la falta de Haspin en cromosomas metafásicos provoca la separación de las cromátidas hermanas, como ya estaba descrito para vertebrados (Dai et al., 2006). Pds5 tiene dos papeles en la regulación de la dinámica de las cohesinas, por un lado colabora junto con Eco en el mantenimiento de las cohesinas unidas y, por otro, también participa en su liberación en colaboración con wapl (Dorsett, 2009). Esto, junto con el hecho de que pds5 determina la posición de Haspin (Yamagishi et al., 2010), hace que se plantee un mecanismo de acción en el que pds5 localiza Haspin en el DNA y ambas actúan sobre las cohesinas previamente unidas al DNA. Se ha publicado recientemente que la falta de cohesinas y CTCF en células de mamífero provoca condensación del DNA interfásico (Tark-Dame et al., 2014), resultado acorde a lo que nosotros observamos en *Drosophila* en ausencia de Haspin y pds5. No hemos investigado nada a cerca de la relación entre Haspin, pds5 y cohesinas a la hora de organizar la estructura de la cromatina, pero sería interesante comprobar si efectivamente las cohesinas afectan a la condensación de la cromatina también en *Drosophila*, y si dependen de Haspin y pds5 para llevarlo a cabo.

No hemos estudiado si la relación entre Haspin y pds5 se realiza a través de H3T3ph. Sería interesante comparar sus perfiles de unión a lo largo de todo el genoma; sin embargo, lamentablemente, no hay datos en modENCODE sobre los sitios en los que pds5 está presente. Se ha visto que tanto las cohesinas como Nipped-B están unidas principalmente, pero no de manera exclusiva, a genes transcripcionalmente activos (Misulovin et al., 2008, Schaaf et al., 2009). Esta localización parece opuesta a la que ocupa H3T3ph. No obstante, es posible que aunque H3T3ph esté mayoritariamente en regiones silenciadas ejerza también una función en zonas activas en interfase, como se desprende por su localización en cromosomas politénicos, siendo en éstas en las que colabore con cohesinas.

Hemos comprobado que tanto pds5 como SMC1 (proteína que forma parte del complejo de las cohesinas) tienen efecto *enhacer-blocking* en los Fabs del complejo Bithorax, pero no sabemos si Haspin, pds5 y las cohesinas tienen una dependencia mutua en esta actividad. En mamíferos CTCF colocaliza con cohesinas y las necesita para realizar su función como *insulator* (Parelho et al., 2008, Rubio et al., 2008, Wendt et al., 2008), pero esta relación no se ha visto en *Drosophila* (Bartkuhn et al., 2009). No obstante, se ha descrito que las cohesinas colocalizan con CP190 (Bartkuhn et al., 2009), por lo que, efectivamente, en *Drosophila* también puede que desempeñen esta actividad *insulator*.

Durante mitosis y meiosis, el complejo de las cohesinas tiene una localización espacio-temporal opuesta al complejo de las condensinas como hemos señalado en el apartado 3.1 de la introducción. El complejo de las condensinas no solo es esencial para la formación de los cromosomas mitóticos.

## **Discusión**

Sus sitios de unión pueden interactuar formando agrupaciones de regiones del genoma que de otra manera estarían dispersas, por lo que también desempeñan un papel en la organización tridimensional de la cromatina en el núcleo en interfase (Frosi and Haering, 2015). De los dos grupos de condensinas, condensinas I y condensinas II, solo el segundo se localiza en el núcleo a lo largo de todo el ciclo celular, por lo que las funciones del complejo en interfase quedan reducidas a él. Ambos grupos tienen las mismas proteínas SMC, y se diferencian en las otras tres, que en las condensinas II son CapG2, CapH2 y CapD3. De entre ellas, existen más evidencias de la función de CapH2. A diferencia de los numerosos estudios que se han hecho sobre condensinas en mitosis, poco se ha investigado en interfase (Hirano, 2005, Wood et al., 2010, Carter and Sjogren, 2012).

Se ha visto que las condensinas están implicadas en la formación de los territorios cromosomales, impiden el normal emparejamiento de las cromátidas hermanas, promueven la condensación del núcleo interfásico y, en consecuencia a esta hipercompactación de los cromosomas unidos a la membrana nuclear, pueden colapsar la envoltura nuclear produciendo fuertes alteraciones (Hartl et al., 2008, Bauer et al., 2012, Joyce et al., 2012, Buster et al., 2013). Comparando estos datos con nuestros resultados, vemos que la sobreexpresión de CapH2 produce los mismos efectos que la ausencia de Haspin. Es cierto que la falta de Haspin no induce unos fenotipos tan drásticos, ya que las invaginaciones que observamos en la periferia nuclear son similares, pero en ningún caso se forman esferas intranucleares de lamina como se ha descrito para CapH2.

Comparando los datos de nuestro ChIP-seq con los de CapH2 se aprecia que regiones que están enriquecidas en H3T3ph no contienen CapH2, y en cambio, zonas donde no se encuentra la histona fosforilada presentan gran cantidad de CapH2 unida. Parece que H3T3ph y CapH2 tienen una localización mutuamente excluyente. Nuestros resultados apuntan a que lo que provoca la acumulación de CapH2 en la cromatina en interfase es la falta de Haspin. Detectamos que existe una interacción genética entre Haspin y las condensinas, pues la condensación que produce la falta de Haspin requiere de la acción de CapH2. No obstante, hasta el momento no tenemos evidencias directas de que Haspin regule la unión de CapH2 al DNA. Para ello, se podría realizar un ChIP-seq que nos permita ver si la unión de CapH2 al genoma y las regiones en las que se encuentra varían en presencia o ausencia de Haspin. El análisis de estos datos nos aportaría información sobre esta posible relación.

El hecho de que Haspin medie sus funciones de manera exclusiva a través de H3T3ph es una hipótesis que debe ser comprobada, porque quizás la kinasa está ejerciendo su acción a través de la fosforilación de otra proteína. Recientemente se ha hecho un estudio en células HeLa estimando las proteínas con más probabilidad de ser fosforiladas por Haspin (Maiolica et al., 2014). Destacan tres grupos: proteínas que regulan la mitosis, otras involucradas en el procesamiento del RNA e histonas y proteínas que intervienen en el control de la expresión génica. Demuestran que macroH2A pudiera ser fosforilada por Haspin en la serina de la posición 137, pero proponen otras proteínas que también tienen alta probabilidad de ser dianas para la kinasa.

De entre ellas, resulta interesante la presencia de proteínas que unen específicamente histonas modificadas transcripcionalmente. En ausencia y presencia de Haspin se observan cambios en la fosforilación de proteínas con cromodominios que reconocen residuos di y trimetilados, tal que HP1, que une H3K9, y Polycomb, que lo hace con H3K27. Estas proteínas controlan la represión epigenética de la cromatina (Taverna et al., 2007). Por otro lado, también señalan como posible diana a Nuclear Protein Ataxia-Telangiectasia (NPAT), cuyo ortólogo en *Drosophila* es *multi sex combs* (mx). Esta proteína se ha incluido en el grupo de Polycomb, aunque no pertenece al PRC1 ni PRC2. Se ha visto que mx actúa sobre la expresión de genes homeóticos (Landais et al., 2014). Acorde a nuestros resultados, H3T3ph colocaliza con la marca realizada por PRC2 y reconocida por proteínas del PRC1. Por otro lado, hemos comprobado que la ausencia de Haspin conlleva una menor unión de Polycomb a algunas de las regiones reguladoras de los genes homeóticos, sin valorar la repercusión que tiene sobre otras proteínas PcG. Es posible que Haspin pueda actuar sobre las proteínas de los complejos represores de Polycomb de dos maneras: una mediante la fosforilación de H3T3 como ya hemos señalado, y también de manera directa fosforilando a estas proteínas.

La dinámica de unión al DNA de las cohesinas está regulada por una intrincada red de modificaciones posttranscripcionales (Rudra and Skibbens, 2013). Principalmente están influidas por la acetilación de diferentes subunidades del complejo así como por diferentes sitios diana dentro de cada subunidad. No obstante, otras modificaciones como SUMOilación, ubiquitinación y fosforilación también influyen en ellas. Además, esto no solo se ha visto para la de las propias proteínas que forman el anillo, sino que también afectan a las responsables de su unión, mantenimiento y eliminación del DNA. Todos los ortólogos de Nipped-B están fosforilados (Villen et al., 2007, Dephoure et al., 2008, Wilson-Grady et al., 2008). Recientemente se ha realizado un estudio para ver la importancia de la fosforilación de SCC2 (proteína ortóloga de Nipped-B en *S. Cerevisiae*) y cómo repercute sobre el complejo y en los procesos celulares en los que interviene. Se ha visto que el control de la presencia o ausencia de los residuos fosforilados es fundamental a la hora de mediar sus acciones, ya que tiene varios sitios que se pueden fosforilar. Los efectos producidos por la proteína fosforilada son opuestos al mantenimiento de las cohesinas en el DNA. Sugieren que es necesario que la proteína esté desfosforilada para participar en el mantenimiento del complejo en el DNA o que la fosforilación de la misma hace que atrape a factores que mantienen las cohesinas unidas a los cromosomas (Woodman et al., 2015). Por tanto, Nipped-B fosforilada participaría, junto con otras proteínas como pds5 y wapl en la eliminación de las cohesinas unidas. Previamente ya hemos señalado que pds5 participa tanto en la eliminación como en el mantenimiento, y que su unión a la cromatina depende de Haspin. No obstante, no hemos estudiado cómo este complejo repercute sobre las cohesinas. El extremo C-terminal de Nipped-B tiene un posible sitio de fosforilación diana para Haspin, arginina-treonina-lisina (R-T-K), sitio de reconocimiento por excelencia para su actividad kinasa (Maiolica et al., 2014), por lo que puede que Haspin lo fosforile para que participe en la descarga de las cohesinas junto con pds5.

Sería interesante estudiar si Haspin fosforila directamente a pds5 a pesar de que no se ha detectado en el fosfoproteoma como posible proteína diana. Se ha descrito un comportamiento

## **Discusión**

diferente respecto a la condensación de cromosomas politénicos así como sobre el mantenimiento y separación de cohesinas del genoma en moscas mutantes de pds5 con diferente tipo de escisión del gen. Sorprendentemente, lo que diferenciaba ambas líneas era que una contenía el dominio C-terminal, justo el dominio que contiene la secuencia R-T-K diana para Haspin. Un phos-tag (técnica que permite visualizar una misma proteína con diferentes grados de fosforilación de manera separada en geles de acrilamida) nos permitiría apreciar si pds5 puede presentarse de forma fosforilada y no fosforilada, y, comparando el resultado en ausencia de Haspin, si la kinasa afecta en ello. No obstante, la apreciación de cambios en esta técnica no supondría una implicación directa de Haspin en este proceso, ya que puede ser generada como consecuencia de una cascada de modificaciones posttranscripcionales.

Slimb media la degradación de CapH2 a través del reconocimiento de una secuencia en su extremo C-terminal, que sirve de diana para su ubiquitinación, y, mediante ella, degradación de la proteína (Buster et al., 2013). Para poder llevar a cabo esta función, es necesario que junto a la secuencia de reconocimiento mencionada la proteína esté fosforilada (Deshaies, 1999, Smelkinson and Kalderon, 2006). Se ha visto que CapH2 está altamente regulada a través de la fosforilación, tanto para su degradación como estamos señalando como para su unión al DNA (Kagami et al., 2014).

Recientemente se ha propuesto que CK1 $\alpha$ , mediante la fosforilación de CapH2, regula su degradación (Nguyen et al., 2015). No obstante, necesita de la colaboración de otra kinasa que medie el contacto con Slimb y permita la ubiquitinación de la condensina, y posteriormente, el paso final de su degradación sería mediado por CK1 $\alpha$ . El dominio C-terminal de 3 de las 4 isoformas de polipéptidos que genera el gen CapH2 tienen una secuencia que puede ser reconocida por Haspin, fosforilando en este caso una serina en vez de una treonina. Sería interesante comprobar si Haspin fosforila este residuo y regula la degradación de la proteína. Se podría estudiar a través de cambios en la cantidad de la condensina; si la fosforilación por parte de Haspin, como hipotetizamos, participa en su degradación, la sobreexpresión de la kinasa implicaría una disminución de la cantidad de CapH2. No obstante, cabe señalar que se ha dicho que la kinasa que colabore con CK1 $\alpha$  debe fosforilar la primera treonina o serina de la secuencia S/T-X-X-S/T (Flotow et al., 1990, Yanagawa et al., 2002). En el sitio que hemos visto como probable para que actúe Haspin, sería Haspin la que fosforilaría la segunda serina y no la primera como se ha propuesto.

Con todos los datos obtenidos durante esta tesis planteamos la hipótesis de que Haspin mantiene la cromatina parcialmente descondensada en zonas silenciadas, siendo necesario para ello la actividad kinasa de la proteína y la fosforilación de H3T3. Esta hipótesis explicaría la colocación de H3T3ph y H3K27me3 a lo largo del genoma y la participación de Haspin en la unión de Polycomb a las regiones reguladoras del gen *Abdominal-B*, sugiriendo que la proteína kinasa sería necesaria para abrir la cromatina de manera que proteínas de PcG puedan unirse al DNA y así mediar su acción silenciadora.

Por otro lado, el hecho de que Haspin sea necesaria para que pds5 se una al DNA, así como que la localización en el genoma de H3T3ph y CapH2 sea mutuamente excluyente indica que la relación Haspin-pds5-cohesinas-condensinas también puede existir en interfase.



# Conclusiones





1. La proteína *insulator* CTCF de *Drosophila melanogaster* es necesaria para la expresión de los genes *Halloween spookier*, *shadow* y *noppera-bo* que regulan la síntesis de ecdisona, así como para la homeostasis de los lípidos en las células donde se produce esta hormona.
2. Las proteínas Ibf1, Ibf2, Pita y CP190 forman parte del mismo complejo proteico.
3. Ibf1 e Ibf2 se necesitan mutuamente para estar presentes en *insulator bodies*, mientras que la localización de CP190 es independiente de estas proteínas.
4. Ibf1 e Ibf2 tienen actividad *enhancer-blocking*.
5. La kinasa Haspin de *Drosophila* tiene actividad *enhancer-blocking*.
6. Haspin influye en la localización nuclear de CP190.
7. Haspin modula la estructura, condensación y organización de la cromatina gracias a su dominio kinasa.
8. En *Drosophila melanogaster*, Haspin es la proteína kinasa que fosforila H3T3ph tanto en interfase como en mitosis.
9. H3T3ph está localizada en heterocromatina pericentromérica y en grandes dominios silenciados, flanqueados por proteínas *insulators*.
10. La unión de Polycomb al complejo Bithorax está regulada por Haspin.
11. Pds5 y SMC1, proteínas del complejo de las cohesinas, tienen actividad *enhancer-blocking*.
12. Haspin es necesaria para la unión de pds5 al DNA en interfase.
13. H3T3ph y CapH2, proteína del complejo de las condensinas II, presentan una localización mutuamente excluyente a lo largo del genoma.



# Bibliografía



- Ahanger, S. H. et al. (2013). Functional sub-division of the *Drosophila* genome via chromatin looping: the emerging importance of CP190. *Nucleus* 4(2): 115-122.
- Akhtar, A. and Gasser, S. M. (2007). The nuclear envelope and transcriptional control. *Nat Rev Genet* 8(7): 507-517.
- Ali, T. et al. (2016). Insulators and domains of gene expression. *Curr Opin Genet Dev* 37: 17-26.
- Aoki, T. et al. (2012). Elba, a novel developmentally regulated chromatin boundary factor is a hetero-tripartite DNA binding complex. *Elife* 1: e00171.
- Arnold, C. D. et al. (2014). Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat Genet* 46(7): 685-692.
- Ball, A. R., Jr. et al. (2014). Mechanisms of cohesin-mediated gene regulation and lessons learned from cohesinopathies. *Biochim Biophys Acta* 1839(3): 191-202.
- Banerji, J. et al. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27(2 Pt 1): 299-308.
- Bantignies, F. et al. (2011). Polycomb-dependent regulatory contacts between distant Hox loci in *Drosophila*. *Cell* 144(2): 214-226.
- Barges, S. et al. (2000). The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. *Development* 127(4): 779-790.
- Barski, A. et al. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129(4): 823-837.
- Bartkuhn, M. et al. (2009). Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J* 28(7): 877-888.
- Bauer, C. R. et al. (2012). Condensin II promotes the formation of chromosome territories by inducing axial compaction of polyploid interphase chromosomes. *PLoS Genet* 8(8): e1002873.
- Baxley, R. M. et al. (2011). The role of the Suppressor of Hairy-wing insulator protein in *Drosophila* oogenesis. *Dev Biol* 356(2): 398-410.
- Beisel, C. et al. (2007). Comparing active and repressed expression states of genes controlled by the Polycomb/Trithorax group proteins. *Proc Natl Acad Sci U S A* 104(42): 16615-16620.
- Bell, A. C. et al. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98(3): 387-396.

## **Bibliografia**

- Bergstrom, R. et al. (2010). Transforming growth factor beta promotes complexes between Smad proteins and the CCCTC-binding factor on the H19 imprinting control region chromatin. *J Biol Chem* 285(26): 19727-19737.
- Berry, S. et al. (2015). Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. *Elife* 4.
- Bhaumik, S. R. et al. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14(11): 1008-1016.
- Bohla, D. et al. (2014). A functional insulator screen identifies NURF and dREAM components to be required for enhancer-blocking. *PLoS One* 9(9): e107765.
- Boulet, A. M. et al. (1991). Molecular definition of the morphogenetic and regulatory functions and the cis-regulatory elements of the *Drosophila* Abd-B homeotic gene. *Development* 111(2): 393-405.
- Boyarchuk, Y. et al. (2007). Bub1 is essential for assembly of the functional inner centromere. *J Cell Biol* 176(7): 919-928.
- Boyle, S. et al. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet* 10(3): 211-219.
- Bradbury, E. M. (1992). Reversible histone modifications and the chromosome cell cycle. *Bioessays* 14(1): 9-16.
- Bushey, A. M. et al. (2009). Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev* 23(11): 1338-1350.
- Buster, D. W. et al. (2013). SCF<sup>Slimb</sup> ubiquitin ligase suppresses condensin II-mediated nuclear reorganization by degrading Cap-H2. *J Cell Biol* 201(1): 49-63.
- Buszczak, M. et al. (1999). Ecdysone response genes govern egg chamber development during mid-oogenesis in *Drosophila*. *Development* 126(20): 4581-4589.
- Butcher, R. D. et al. (2004). The *Drosophila* centrosome-associated protein CP190 is essential for viability but not for cell division. *J Cell Sci* 117(Pt 7): 1191-1199.
- Butler, J. E. and Kadonaga, J. T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev* 16(20): 2583-2592.
- Byrd, K. and Corces, V. G. (2003). Visualization of chromatin domains created by the gypsy insulator of *Drosophila*. *J Cell Biol* 162(4): 565-574.
- Caldwell, P. E. et al. (2005). Ras activity in the *Drosophila* prothoracic gland regulates body size and developmental rate via ecdysone release. *Curr Biol* 15(20): 1785-1795.
- Cao, R. and Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* 14(2): 155-164.

- Carmena, M. et al. (2012). The chromosomal passenger complex activates Polo kinase at centromeres. *PLoS Biol* 10(1): e1001250.
- Carney, G. E. and Bender, M. (2000). The *Drosophila* ecdysone receptor (EcR) gene is required maternally for normal oogenesis. *Genetics* 154(3): 1203-1211.
- Carretero, M. et al. (2013). Pds5B is required for cohesion establishment and Aurora B accumulation at centromeres. *EMBO J* 32(22): 2938-2949.
- Carter, S. D. and Sjogren, C. (2012). The SMC complexes, DNA and chromosome topology: right or knot? *Crit Rev Biochem Mol Biol* 47(1): 1-16.
- Casolari, J. M. et al. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* 117(4): 427-439.
- Cavalli, G. and Paro, R. (1998). The *Drosophila* Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. *Cell* 93(4): 505-518.
- Celniker, S. E. et al. (1989). The molecular genetics of the bithorax complex of *Drosophila*: characterization of the products of the Abdominal-B domain. *Genes Dev* 3(9): 1424-1436.
- Ciavatta, D. et al. (2007). *Drosophila* CTCF is required for Fab-8 enhancer blocking activity in S2 cells. *J Mol Biol* 373(2): 233-239.
- Clark, A. J. and Block, K. (1959). The absence of sterol synthesis in insects. *J Biol Chem* 234: 2578-2582.
- Clarkson, M. and Saint, R. (1999). A His2AvDGFP fusion gene complements a lethal His2AvD mutant allele and provides an in vivo marker for *Drosophila* chromosome behavior. *DNA Cell Biol* 18(6): 457-462.
- Cleard, F. et al. (2006). Probing long-distance regulatory interactions in the *Drosophila melanogaster* bithorax complex using Dam identification. *Nat Genet* 38(8): 931-935.
- Cobbe, N. et al. (2006). Diverse mitotic and interphase functions of condensins in *Drosophila*. *Genetics* 172(2): 991-1008.
- Colombani, J. et al. (2005). Antagonistic actions of ecdysone and insulins determine final size in *Drosophila*. *Science* 310(5748): 667-670.
- Comet, I. et al. (2006). PRE-mediated bypass of two Su(Hw) insulators targets PcG proteins to a downstream promoter. *Dev Cell* 11(1): 117-124.
- Comet, I. et al. (2011). A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc Natl Acad Sci U S A* 108(6): 2294-2299.
- Corces, V. G. (1995). Chromatin insulators. Keeping enhancers under control. *Nature* 376(6540): 462-463.



## **Bibliografia**

Crane, E. et al. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* 523(7559): 240-244.

Cremer, T. and Cremer, M. (2010). Chromosome territories. *Cold Spring Harb Perspect Biol* 2(3): a003889.

Cuartero, S. et al. (2014). Ibf1 and Ibf2 are novel CP190-interacting proteins required for insulator function. *EMBO J* 33(6): 637-647.

Cubenas-Potts, C. and Corces, V. G. (2015). Architectural proteins, transcription, and the three-dimensional organization of the genome. *FEBS Lett* 589(20 Pt A): 2923-2930.

Chavez, V. M. et al. (2000). The *Drosophila* disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels. *Development* 127(19): 4115-4126.

Chen, Q. et al. (2005). Multiple Promoter Targeting Sequences exist in Abdominal-B to regulate long-range gene activation. *Dev Biol* 286(2): 629-636.

Chung, J. H. et al. (1993). A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* 74(3): 505-514.

Dai, J. et al. (2005). The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev* 19(4): 472-488.

Dai, J. et al. (2006). Regulation of mitotic chromosome cohesion by Haspin and Aurora B. *Dev Cell* 11(5): 741-750.

Danielsen, E. T. et al. (2014). Transcriptional control of steroid biosynthesis genes in the *Drosophila* prothoracic gland by ventral veins lacking and knirps. *PLoS Genet* 10(6): e1004343.

de Graaf, C. A. and van Steensel, B. (2013). Chromatin organization: form to function. *Curr Opin Genet Dev* 23(2): 185-190.

Dejardin, J. et al. (2005). Recruitment of *Drosophila* Polycomb group proteins to chromatin by DSP1. *Nature* 434(7032): 533-538.

Dekker, J. et al. (2002). Capturing chromosome conformation. *Science* 295(5558): 1306-1311.

Deng, H. et al. (2008). Ectopic histone H3S10 phosphorylation causes chromatin structure remodeling in *Drosophila*. *Development* 135(4): 699-705.

Deng, H. and Kerppola, T. K. (2013). Regulation of *Drosophila* metamorphosis by xenobiotic response regulators. *PLoS Genet* 9(2): e1003263.

Denissov, S. et al. (2014). Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* 141(3): 526-537.

- Dephoure, N. et al. (2008). A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A* 105(31): 10762-10767.
- Deshaies, R. J. (1999). SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol* 15: 435-467.
- Di Croce, L. and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nat Struct Mol Biol* 20(10): 1147-1155.
- Dixon, J. R. et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485(7398): 376-380.
- Donze, D. et al. (1999). The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*. *Genes Dev* 13(6): 698-708.
- Dorsett, D. (2009). Cohesin, gene expression and development: lessons from *Drosophila*. *Chromosome Res* 17(2): 185-200.
- Dorsett, D. et al. (2005). Effects of sister chromatid cohesion proteins on cut gene expression during wing development in *Drosophila*. *Development* 132(21): 4743-4753.
- Downen, J. M. et al. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159(2): 374-387.
- Doyle, B. et al. (2014). Chromatin loops as allosteric modulators of enhancer-promoter interactions. *PLoS Comput Biol* 10(10): e1003867.
- Duncan, I. M. (1982). Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of *Drosophila melanogaster*. *Genetics* 102(1): 49-70.
- Ebert, A. et al. (2004). Su(var) genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev* 18(23): 2973-2983.
- Eldholm, V. et al. (2014). CTCF mediates the TERT enhancer-promoter interactions in lung cancer cells: identification of a novel enhancer region involved in the regulation of TERT gene. *Int J Cancer* 134(10): 2305-2313.
- Enya, S. et al. (2014). A Halloween gene noppera-bo encodes a glutathione S-transferase essential for ecdysteroid biosynthesis via regulating the behaviour of cholesterol in *Drosophila*. *Sci Rep* 4: 6586.
- Ephrussi, B. and Herold, J. L. (1944). Studies of Eye Pigments of *Drosophila*. I. Methods of Extraction and Quantitative Estimation of the Pigment Components. *Genetics* 29(2): 148-175.
- Erokhin, M. et al. (2011). Insulators form gene loops by interacting with promoters in *Drosophila*. *Development* 138(18): 4097-4106.
- Erokhin, M. et al. (2015). Eukaryotic enhancers: common features, regulation, and participation in diseases. *Cell Mol Life Sci* 72(12): 2361-2375.

## ***Bibliografia***

---

- Eswaran, J. et al. (2009). Structure and functional characterization of the atypical human kinase haspin. *Proc Natl Acad Sci U S A* 106(48): 20198-20203.
- Fasulo, B. et al. (2012). The Drosophila MI-2 chromatin-remodeling factor regulates higher-order chromatin structure and cohesin dynamics in vivo. *PLoS Genet* 8(8): e1002878.
- Filion, G. J. et al. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* 143(2): 212-224.
- Finlan, L. E. et al. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet* 4(3): e1000039.
- Flotow, H. et al. (1990). Phosphate groups as substrate determinants for casein kinase I action. *J Biol Chem* 265(24): 14264-14269.
- Font-Burgada, J. et al. (2014). chroGPS, a global chromatin positioning system for the functional analysis and visualization of the epigenome. *Nucleic Acids Res* 42(4): 2126-2137.
- Frosi, Y. and Haering, C. H. (2015). Control of chromosome interactions by condensin complexes. *Curr Opin Cell Biol* 34: 94-100.
- Gao, J. et al. (2007). Functional role of CCCTC binding factor in insulin-stimulated cell proliferation. *Cell Prolif* 40(6): 795-808.
- Garen, A. et al. (1977). Roles of ecdysone in Drosophila development. *Proc Natl Acad Sci U S A* 74(11): 5099-5103.
- Garske, A. L. et al. (2010). Combinatorial profiling of chromatin binding modules reveals multisite discrimination. *Nat Chem Biol* 6(4): 283-290.
- Gaszner, M. et al. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev* 13(16): 2098-2107.
- Geisler, S. J. and Paro, R. (2015). Trithorax and Polycomb group-dependent regulation: a tale of opposing activities. *Development* 142(17): 2876-2887.
- Gentleman, R. C. et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5(10): R80.
- Gerasimova, T. I. et al. (2000). A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* 6(5): 1025-1035.
- Gerasimova, T. I. and Corces, V. G. (1998). Polycomb and trithorax group proteins mediate the function of a chromatin insulator. *Cell* 92(4): 511-521.
- Gerasimova, T. I. and Corces, V. G. (2001). Chromatin insulators and boundaries: effects on transcription and nuclear organization. *Annu Rev Genet* 35: 193-208.

- Gerasimova, T. I. et al. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Mol Cell* 28(5): 761-772.
- Geyer, P. K. and Corces, V. G. (1992). DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev* 6(10): 1865-1873.
- Ghenoiu, C. et al. (2013). Autoinhibition and Polo-dependent multisite phosphorylation restrict activity of the histone H3 kinase Haspin to mitosis. *Mol Cell* 52(5): 734-745.
- Gibbens, Y. Y. et al. (2011). Neuroendocrine regulation of *Drosophila* metamorphosis requires TGFbeta/Activin signaling. *Development* 138(13): 2693-2703.
- Gilbert, L. I. (2004). Halloween genes encode P450 enzymes that mediate steroid hormone biosynthesis in *Drosophila melanogaster*. *Mol Cell Endocrinol* 215(1-2): 1-10.
- Gilbert, L. I. et al. (2002). Control and biochemical nature of the ecdysteroidogenic pathway. *Annu Rev Entomol* 47: 883-916.
- Golovnin, A. et al. (2008). 'Insulator bodies' are aggregates of proteins but not of insulators. *EMBO Rep* 9(5): 440-445.
- Gomos-Klein, J. et al. (2007). CTCF-independent, but not CTCF-dependent, elements significantly contribute to TCR-alpha locus control region activity. *J Immunol* 179(2): 1088-1095.
- Goto, H. et al. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells* 7(1): 11-17.
- Groschel, S. et al. (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* 157(2): 369-381.
- Gruzdeva, N. et al. (2005). The Mcp element from the bithorax complex contains an insulator that is capable of pairwise interactions and can facilitate enhancer-promoter communication. *Mol Cell Biol* 25(9): 3682-3689.
- Gurudatta, B. V. and Corces, V. G. (2009). Chromatin insulators: lessons from the fly. *Brief Funct Genomic Proteomic* 8(4): 276-282.
- Hadjur, S. et al. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460(7253): 410-413.
- Hagstrom, K. et al. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. *Genes Dev* 10(24): 3202-3215.
- Hallson, G. et al. (2008). The *Drosophila* cohesin subunit Rad21 is a trithorax group (trxG) protein. *Proc Natl Acad Sci U S A* 105(34): 12405-12410.
- Hart, C. M. et al. (1997). The scs' boundary element: characterization of boundary element-associated factors. *Mol Cell Biol* 17(2): 999-1009.

## **Bibliografia**

Hartl, T. A. et al. (2008). Chromosome alignment and transvection are antagonized by condensin II. *Science* 322(5906): 1384-1387.

Heger, P. et al. (2012). The chromatin insulator CTCF and the emergence of metazoan diversity. *Proc Natl Acad Sci U S A* 109(43): 17507-17512.

Heger, P. et al. (2009). Loss of the insulator protein CTCF during nematode evolution. *BMC Mol Biol* 10: 84.

Heger, P. and Wiehe, T. (2014). New tools in the box: an evolutionary synopsis of chromatin insulators. *Trends Genet* 30(5): 161-171.

Hendzel, M. J. et al. (1997). Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* 106(6): 348-360.

Herold, M. et al. (2012). CTCF: insights into insulator function during development. *Development* 139(6): 1045-1057.

Hietakangas, V. and Cohen, S. M. (2009). Regulation of tissue growth through nutrient sensing. *Annu Rev Genet* 43: 389-410.

Higgins, J. M. (2001). Haspin-like proteins: a new family of evolutionarily conserved putative eukaryotic protein kinases. *Protein Sci* 10(8): 1677-1684.

Higgins, J. M. (2010). Haspin: a newly discovered regulator of mitotic chromosome behavior. *Chromosoma* 119(2): 137-147.

Hirano, T. (2005). Condensins: organizing and segregating the genome. *Curr Biol* 15(7): R265-275.

Holohan, E. E. et al. (2007). CTCF genomic binding sites in *Drosophila* and the organisation of the bithorax complex. *PLoS Genet* 3(7): e112.

Hou, C. et al. (2010). Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci U S A* 107(8): 3651-3656.

Hou, C. et al. (2012). Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell* 48(3): 471-484.

Hou, C. et al. (2008). CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A* 105(51): 20398-20403.

Huang, X. et al. (2005). A *Drosophila* model of the Niemann-Pick type C lysosome storage disease: *dnpc1a* is required for molting and sterol homeostasis. *Development* 132(22): 5115-5124.

Iampietro, C. et al. (2010). Initiator elements function to determine the activity state of BX-C enhancers. *PLoS Genet* 6(12): e1001260.

- Ishihara, K. et al. (2006). CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell* 23(5): 733-742.
- Ishii, K. et al. (2002). Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* 109(5): 551-562.
- Iwasaki, O. et al. (2010). Centromeric localization of dispersed Pol III genes in fission yeast. *Mol Biol Cell* 21(2): 254-265.
- Jans, J. et al. (2009). A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. *Genes Dev* 23(5): 602-618.
- Jeong, M. W. et al. (2013). Mitogen-activated protein kinase phosphatase 2 regulates histone H3 phosphorylation via interaction with vaccinia-related kinase 1. *Mol Biol Cell* 24(3): 373-384.
- Jin, Y. et al. (1999). JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in *Drosophila*. *Mol Cell* 4(1): 129-135.
- Joyce, E. F. et al. (2012). Identification of genes that promote or antagonize somatic homolog pairing using a high-throughput FISH-based screen. *PLoS Genet* 8(5): e1002667.
- Kadonaga, J. T. (2012). Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip Rev Dev Biol* 1(1): 40-51.
- Kagami, Y. et al. (2014). Mps1 phosphorylation of condensin II controls chromosome condensation at the onset of mitosis. *J Cell Biol* 205(6): 781-790.
- Kagey, M. H. et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467(7314): 430-435.
- Kamoshida, Y. et al. (2012). Ecdysone receptor (EcR) suppresses lipid accumulation in the *Drosophila* fat body via transcription control. *Biochem Biophys Res Commun* 421(2): 203-207.
- Kang, T. H. et al. (2007). Mitotic histone H3 phosphorylation by vaccinia-related kinase 1 in mammalian cells. *Mol Cell Biol* 27(24): 8533-8546.
- Karch, F. et al. (1994). Mcp and Fab-7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of *Drosophila melanogaster*. *Nucleic Acids Res* 22(15): 3138-3146.
- Kaufman, T. C. et al. (1980). Cytogenetic Analysis of Chromosome 3 in *DROSOPHILA MELANOGASTER*: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. *Genetics* 94(1): 115-133.
- Kehle, J. et al. (1998). dMi-2, a hunchback-interacting protein that functions in polycomb repression. *Science* 282(5395): 1897-1900.
- Keightley, D. A. et al. (1990). Involvement of translation and transcription in insect steroidogenesis. *Mol Cell Endocrinol* 74(3): 229-237.

## **Bibliografia**

- Kellogg, D. R. and Alberts, B. M. (1992). Purification of a multiprotein complex containing centrosomal proteins from the *Drosophila* embryo by chromatography with low-affinity polyclonal antibodies. *Mol Biol Cell* 3(1): 1-11.
- Kellum, R. and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64(5): 941-950.
- Kellum, R. and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol Cell Biol* 12(5): 2424-2431.
- Kennison, J. A. and Tamkun, J. W. (1988). Dosage-dependent modifiers of polycomb and antennapedia mutations in *Drosophila*. *Proc Natl Acad Sci U S A* 85(21): 8136-8140.
- Khadjeh, S. et al. (2012). Divergent role of the Hox gene *Antennapedia* in spiders is responsible for the convergent evolution of abdominal limb repression. *Proc Natl Acad Sci U S A* 109(13): 4921-4926.
- Kharchenko, P. V. et al. (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471(7339): 480-485.
- Kingston, R. E. and Tamkun, J. W. (2014). Transcriptional regulation by trithorax-group proteins. *Cold Spring Harb Perspect Biol* 6(10): a019349.
- Kirkland, J. G. et al. (2013). TFIIIC bound DNA elements in nuclear organization and insulation. *Biochim Biophys Acta* 1829(3-4): 418-424.
- Klymenko, T. and Muller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep* 5(4): 373-377.
- Koshland, D. and Strunnikov, A. (1996). Mitotic chromosome condensation. *Annu Rev Cell Dev Biol* 12: 305-333.
- Kostyuchenko, M. et al. (2009). Zeste can facilitate long-range enhancer-promoter communication and insulator bypass in *Drosophila melanogaster*. *Chromosoma* 118(5): 665-674.
- Kyrchanova, O. et al. (2013). Effective blocking of the white enhancer requires cooperation between two main mechanisms suggested for the insulator function. *PLoS Genet* 9(7): e1003606.
- Kyrchanova, O. et al. (2008). Functional interaction between the Fab-7 and Fab-8 boundaries and the upstream promoter region in the *Drosophila* Abd-B gene. *Mol Cell Biol* 28(12): 4188-4195.
- Labrador, M. and Corces, V. G. (2002). Setting the boundaries of chromatin domains and nuclear organization. *Cell* 111(2): 151-154.
- Landais, S. et al. (2014). Persistent replicative stress alters polycomb phenotypes and tissue homeostasis in *Drosophila melanogaster*. *Cell Rep* 7(3): 859-870.
- Langmead, B. et al. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3): R25.

- Lanzuolo, C. et al. (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* 9(10): 1167-1174.
- Larschan, E. et al. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol Cell* 28(1): 121-133.
- Laundrie, B. et al. (2003). Germline cell death is inhibited by P-element insertions disrupting the dcp-1/pita nested gene pair in *Drosophila*. *Genetics* 165(4): 1881-1888.
- Lawrence, M. et al. (2009). rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics* 25(14): 1841-1842.
- Layalle, S. et al. (2008). The TOR pathway couples nutrition and developmental timing in *Drosophila*. *Dev Cell* 15(4): 568-577.
- Le Dily, F. et al. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev* 28(19): 2151-2162.
- Lee, J. et al. (2011). Condensins I and II are essential for construction of bivalent chromosomes in mouse oocytes. *Mol Biol Cell* 22(18): 3465-3477.
- Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* 276(5688): 565-570.
- Li, H. B. et al. (2011). Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in *Drosophila melanogaster*. *Mol Cell Biol* 31(4): 616-625.
- Li, L. et al. (2015). Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* 58(2): 216-231.
- Li, Q. and Stamatoyannopoulos, G. (1994). Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. *Blood* 84(5): 1399-1401.
- Li, X. and Noll, M. (1994). Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the *Drosophila* embryo. *EMBO J* 13(2): 400-406.
- Li, X. et al. (2004). Direct association with inner centromere protein (INCENP) activates the novel chromosomal passenger protein, Aurora-C. *J Biol Chem* 279(45): 47201-47211.
- Lin, Q. et al. (2010). Chromatin insulator and the promoter targeting sequence modulate the timing of long-range enhancer-promoter interactions in the *Drosophila* embryo. *Dev Biol* 339(2): 329-337.
- Liscum, L. and Faust, J. R. (1987). Low density lipoprotein (LDL)-mediated suppression of cholesterol synthesis and LDL uptake is defective in Niemann-Pick type C fibroblasts. *J Biol Chem* 262(35): 17002-17008.



## **Bibliografia**

- Lobanenkov, V. V. et al. (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* 5(12): 1743-1753.
- Luger, K. et al. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389(6648): 251-260.
- Lunyak, V. V. et al. (2007). Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* 317(5835): 248-251.
- Lupianez, D. G. et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161(5): 1012-1025.
- Lupo, R. et al. (2001). Drosophila chromosome condensation proteins Topoisomerase II and Barren colocalize with Polycomb and maintain Fab-7 PRE silencing. *Mol Cell* 7(1): 127-136.
- Maeda, R. K. and Karch, F. (2006). The ABC of the BX-C: the bithorax complex explained. *Development* 133(8): 1413-1422.
- Mahadevan, L. C. et al. (1991). Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. *Cell* 65(5): 775-783.
- Maiolica, A. et al. (2014). Modulation of the chromatin phosphoproteome by the Haspin protein kinase. *Mol Cell Proteomics* 13(7): 1724-1740.
- Maksimenko, O. et al. (2015). Two new insulator proteins, Pita and ZIPIC, target CP190 to chromatin. *Genome Res* 25(1): 89-99.
- Mallin, D. R. et al. (1998). Polycomb group repression is blocked by the Drosophila suppressor of Hairy-wing [su(Hw)] insulator. *Genetics* 148(1): 331-339.
- Markaki, Y. et al. (2009). Phosphorylation of histone H3 at Thr3 is part of a combinatorial pattern that marks and configures mitotic chromatin. *J Cell Sci* 122(Pt 16): 2809-2819.
- Marshall, W. F. (2002). Order and disorder in the nucleus. *Curr Biol* 12(5): R185-192.
- Matharu, N. K. et al. (2010). Vertebrate homologue of Drosophila GAGA factor. *J Mol Biol* 400(3): 434-447.
- McBrayer, Z. et al. (2007). Prothoracicotropic hormone regulates developmental timing and body size in Drosophila. *Dev Cell* 13(6): 857-871.
- Mendjan, S. et al. (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. *Mol Cell* 21(6): 811-823.
- Mifsud, B. et al. (2015). Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet* 47(6): 598-606.

- Mihaly, J. et al. (2006). Dissecting the regulatory landscape of the Abd-B gene of the bithorax complex. *Development* 133(15): 2983-2993.
- Mirth, C. et al. (2005). The role of the prothoracic gland in determining critical weight for metamorphosis in *Drosophila melanogaster*. *Curr Biol* 15(20): 1796-1807.
- Misulovin, Z. et al. (2008). Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome. *Chromosoma* 117(1): 89-102.
- Moeller, M. E. et al. (2013). Dynamic feedback circuits function as a switch for shaping a maturation-inducing steroid pulse in *Drosophila*. *Development* 140(23): 4730-4739.
- Mohan, M. et al. (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J* 26(19): 4203-4214.
- Moon, H. et al. (2005). CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep* 6(2): 165-170.
- Moreau, P. et al. (1981). The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res* 9(22): 6047-6068.
- Muller, J. and Kassis, J. A. (2006). Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. *Curr Opin Genet Dev* 16(5): 476-484.
- Muller, M. et al. (1999). The mcp element from the *Drosophila melanogaster* bithorax complex mediates long-distance regulatory interactions. *Genetics* 153(3): 1333-1356.
- Muravyova, E. et al. (2001). Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science* 291(5503): 495-498.
- Murawska, M. et al. (2011). Stress-induced PARP activation mediates recruitment of *Drosophila* Mi-2 to promote heat shock gene expression. *PLoS Genet* 7(7): e1002206.
- Nativio, R. et al. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet* 5(11): e1000739.
- Naumova, N. et al. (2013). Organization of the mitotic chromosome. *Science* 342(6161): 948-953.
- Negre, N. et al. (2010). A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* 6(1): e1000814.
- Nguyen, A. L. et al. (2014). Phosphorylation of threonine 3 on histone H3 by haspin kinase is required for meiosis I in mouse oocytes. *J Cell Sci* 127(23): 5066-5078.
- Nguyen, A. L. et al. (2014). Phosphorylation of threonine 3 on histone H3 by haspin kinase is required for meiosis I in mouse oocytes. *J Cell Sci* 127(Pt 23): 5066-5078.
- Nguyen, H. Q. et al. (2015). *Drosophila* casein kinase I alpha regulates homolog pairing and genome organization by modulating condensin II subunit Cap-H2 levels. *PLoS Genet* 11(2): e1005014.

## **Bibliografia**

Niwa, R. et al. (2004). CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid biosynthesis in the prothoracic glands of Bombyx and Drosophila. *J Biol Chem* 279(34): 35942-35949.

Niwa, R. et al. (2010). Non-molting glossy/shroud encodes a short-chain dehydrogenase/reductase that functions in the 'Black Box' of the ecdysteroid biosynthesis pathway. *Development* 137(12): 1991-1999.

Niwa, R. and Niwa, Y. S. (2014). Enzymes for ecdysteroid biosynthesis: their biological functions in insects and beyond. *Biosci Biotechnol Biochem* 78(8): 1283-1292.

Noordermeer, D. et al. (2011). The dynamic architecture of Hox gene clusters. *Science* 334(6053): 222-225.

Nora, E. P. et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485(7398): 381-385.

Oegema, K. et al. (1995). The cell cycle-dependent localization of the CP190 centrosomal protein is determined by the coordinate action of two separable domains. *J Cell Biol* 131(5): 1261-1273.

Ohtsuki, S. and Levine, M. (1998). GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. *Genes Dev* 12(21): 3325-3330.

Oktaba, K. et al. (2008). Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. *Dev Cell* 15(6): 877-889.

Oldham, S. and Hafen, E. (2003). Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. *Trends Cell Biol* 13(2): 79-85.

Ong, C. T. and Corces, V. G. (2014). CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet* 15(4): 234-246.

Ono, H. (2014). Ecdysone differentially regulates metamorphic timing relative to 20-hydroxyecdysone by antagonizing juvenile hormone in Drosophila melanogaster. *Dev Biol* 391(1): 32-42.

Ono, H. et al. (2006). Spook and Spookier code for stage-specific components of the ecdysone biosynthetic pathway in Diptera. *Dev Biol* 298(2): 555-570.

Ono, T. et al. (2004). Spatial and temporal regulation of Condensins I and II in mitotic chromosome assembly in human cells. *Mol Biol Cell* 15(7): 3296-3308.

Oulion, S. et al. (2011). Evolution of repeated structures along the body axis of jawed vertebrates, insights from the Scyliorhinus canicula Hox code. *Evol Dev* 13(3): 247-259.

Page, A. R. et al. (2005). Spotted-dick, a zinc-finger protein of Drosophila required for expression of Orc4 and S phase. *EMBO J* 24(24): 4304-4315.

Pai, C. Y. et al. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol Cell* 16(5): 737-748.

- Palstra, R. J. and Grosveld, F. (2012). Transcription factor binding at enhancers: shaping a genomic regulatory landscape in flux. *Front Genet* 3: 195.
- Pankotai, T. et al. (2010). Genes of the ecdysone biosynthesis pathway are regulated by the dATAC histone acetyltransferase complex in *Drosophila*. *Mol Cell Biol* 30(17): 4254-4266.
- Papp, B. and Muller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev* 20(15): 2041-2054.
- Parelho, V. et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132(3): 422-433.
- Parkhurst, S. M. et al. (1988). The *Drosophila* su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. *Genes Dev* 2(10): 1205-1215.
- Parvy, J. P. et al. (2005). A role for betaFTZ-F1 in regulating ecdysteroid titers during post-embryonic development in *Drosophila melanogaster*. *Dev Biol* 282(1): 84-94.
- Pasini, D. et al. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* 27(10): 3769-3779.
- Perez-Lluch, S. et al. (2008). Characterization of new regulatory elements within the *Drosophila* bithorax complex. *Nucleic Acids Res* 36(21): 6926-6933.
- Peters, J. M. et al. (2008). The cohesin complex and its roles in chromosome biology. *Genes Dev* 22(22): 3089-3114.
- Petryk, A. et al. (2003). Shade is the *Drosophila* P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect molting hormone 20-hydroxyecdysone. *Proc Natl Acad Sci U S A* 100(24): 13773-13778.
- Phillips-Cremins, J. E. et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153(6): 1281-1295.
- Phillips, J. E. and Corces, V. G. (2009). CTCF: master weaver of the genome. *Cell* 137(7): 1194-1211.
- Pickersgill, H. et al. (2006). Characterization of the *Drosophila melanogaster* genome at the nuclear lamina. *Nat Genet* 38(9): 1005-1014.
- Polioudaki, H. et al. (2004). Mitotic phosphorylation of histone H3 at threonine 3. *FEBS Lett* 560(1-3): 39-44.
- Pope, B. D. et al. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature* 515(7527): 402-405.
- Qian, J. et al. (2013). Aurora B defines its own chromosomal targeting by opposing the recruitment of the phosphatase scaffold Repo-Man. *Curr Biol* 23(12): 1136-1143.

## **Bibliografia**

- Qian, J. et al. (2011). PP1/Repo-man dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting. *Curr Biol* 21(9): 766-773.
- Rao, S. S. et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159(7): 1665-1680.
- Rea, S. et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406(6796): 593-599.
- Recillas-Targa, F. et al. (2002). Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A* 99(10): 6883-6888.
- Reddy, K. L. et al. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452(7184): 243-247.
- Regnard, C. et al. (2011). Global analysis of the relationship between JIL-1 kinase and transcription. *PLoS Genet* 7(3): e1001327.
- Rewitz, K. F. et al. (2009). The insect neuropeptide PTTH activates receptor tyrosine kinase torso to initiate metamorphosis. *Science* 326(5958): 1403-1405.
- Ringrose, L. and Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* 134(2): 223-232.
- Robinson, P. J. and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* 16(3): 336-343.
- Rollins, R. A. et al. (2004). Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. *Mol Cell Biol* 24(8): 3100-3111.
- Rollins, R. A. et al. (1999). Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics* 152(2): 577-593.
- Roseman, R. R. et al. (1993). The su(Hw) protein insulates expression of the Drosophila melanogaster white gene from chromosomal position-effects. *EMBO J* 12(2): 435-442.
- Rubio, E. D. et al. (2008). CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* 105(24): 8309-8314.
- Rudra, S. and Skibbens, R. V. (2013). Cohesin codes - interpreting chromatin architecture and the many facets of cohesin function. *J Cell Sci* 126(Pt 1): 31-41.
- Sanyal, A. et al. (2012). The long-range interaction landscape of gene promoters. *Nature* 489(7414): 109-113.
- Sasai, K. et al. (2004). Aurora-C kinase is a novel chromosomal passenger protein that can complement Aurora-B kinase function in mitotic cells. *Cell Motil Cytoskeleton* 59(4): 249-263.

- Sawado, T. et al. (2003). The beta -globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation. *Genes Dev* 17(8): 1009-1018.
- Schaaf, C. A. et al. (2009). Regulation of the Drosophila Enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. *PLoS One* 4(7): e6202.
- Schmitges, F. W. et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol Cell* 42(3): 330-341.
- Schoborg, T. et al. (2013). Chromatin insulator bodies are nuclear structures that form in response to osmotic stress and cell death. *J Cell Biol* 202(2): 261-276.
- Schwartz, M. B. et al. (1984). Analysis of metamorphosis in *Drosophila melanogaster*: characterization of giant, an ecdysteroid-deficient mutant. *Dev Biol* 103(1): 85-95.
- Schwartz, Y. B. et al. (2012). Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Res* 22(11): 2188-2198.
- Seitan, V. C. et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res* 23(12): 2066-2077.
- Seitan, V. C. and Merkenschlager, M. (2012). Cohesin and chromatin organisation. *Curr Opin Genet Dev* 22(2): 93-100.
- Sekimata, M. et al. (2009). CCCTC-binding factor and the transcription factor T-bet orchestrate T helper 1 cell-specific structure and function at the interferon-gamma locus. *Immunity* 31(4): 551-564.
- Sewer, M. B. and Waterman, M. R. (2003). ACTH modulation of transcription factors responsible for steroid hydroxylase gene expression in the adrenal cortex. *Microsc Res Tech* 61(3): 300-307.
- Sexton, T. and Cavalli, G. (2015). The role of chromosome domains in shaping the functional genome. *Cell* 160(6): 1049-1059.
- Sexton, T. et al. (2007). Gene regulation through nuclear organization. *Nat Struct Mol Biol* 14(11): 1049-1055.
- Sexton, T. et al. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148(3): 458-472.
- Shahbazian, M. D. and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* 76: 75-100.
- Sheikh, B. N. et al. (2015). MOZ and BMI1 play opposing roles during Hox gene activation in ES cells and in body segment identity specification in vivo. *Proc Natl Acad Sci U S A* 112(17): 5437-5442.
- Shimell, M. J. et al. (1994). Enhancer point mutation results in a homeotic transformation in *Drosophila*. *Science* 264(5161): 968-971.

## **Bibliografia**

Sigrist, C. J. and Pirrotta, V. (1997). Chromatin insulator elements block the silencing of a target gene by the *Drosophila* polycomb response element (PRE) but allow trans interactions between PREs on different chromosomes. *Genetics* 147(1): 209-221.

Sipos, L. and Gyurkovics, H. (2005). Long-distance interactions between enhancers and promoters. *FEBS J* 272(13): 3253-3259.

Sliter, T. J. and Gilbert, L. I. (1992). Developmental arrest and ecdysteroid deficiency resulting from mutations at the *dre4* locus of *Drosophila*. *Genetics* 130(3): 555-568.

Smelkinson, M. G. and Kalderon, D. (2006). Processing of the *Drosophila* hedgehog signaling effector Ci-155 to the repressor Ci-75 is mediated by direct binding to the SCF component Slimb. *Curr Biol* 16(1): 110-116.

Smith, H. F. et al. (2013). Maintenance of interphase chromosome compaction and homolog pairing in *Drosophila* is regulated by the condensin cap-h2 and its partner Mrg15. *Genetics* 195(1): 127-146.

Sofueva, S. et al. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J* 32(24): 3119-3129.

Soshnikova, N. and Duboule, D. (2009). Epigenetic temporal control of mouse Hox genes in vivo. *Science* 324(5932): 1320-1323.

Splinter, E. et al. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 20(17): 2349-2354.

Strahl, B. D. and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403(6765): 41-45.

Strahl, B. D. and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403(6765): 41-45.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev* 12(5): 599-606.

Symmons, O. et al. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Res* 24(3): 390-400.

Szczerbal, I. et al. (2009). The spatial repositioning of adipogenesis genes is correlated with their expression status in a porcine mesenchymal stem cell adipogenesis model system. *Chromosoma* 118(5): 647-663.

Talamillo, A. et al. (2013). Scavenger receptors mediate the role of SUMO and Ftz-f1 in *Drosophila* steroidogenesis. *PLoS Genet* 9(4): e1003473.

Tan, J. et al. (2007). Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev* 21(9): 1050-1063.

- Tanaka, H. et al. (1999). Identification and characterization of a haploid germ cell-specific nuclear protein kinase (Haspin) in spermatid nuclei and its effects on somatic cells. *J Biol Chem* 274(24): 17049-17057.
- Taniura, H. et al. (1995). A chromatin binding site in the tail domain of nuclear lamins that interacts with core histones. *J Cell Biol* 131(1): 33-44.
- Tark-Dame, M. et al. (2014). Depletion of the chromatin looping proteins CTCF and cohesin causes chromatin compaction: insight into chromatin folding by polymer modelling. *PLoS Comput Biol* 10(10): e1003877.
- Taverna, S. D. et al. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 14(11): 1025-1040.
- Tillib, S. et al. (1999). Trithorax- and Polycomb-group response elements within an Ultrabithorax transcription maintenance unit consist of closely situated but separable sequences. *Mol Cell Biol* 19(7): 5189-5202.
- Tolhuis, B. et al. (2011). Interactions among Polycomb domains are guided by chromosome architecture. *PLoS Genet* 7(3): e1001343.
- Tolhuis, B. et al. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10(6): 1453-1465.
- Torras-Llort, M. et al. (2010). A conserved arginine-rich motif within the hypervariable N-domain of *Drosophila* centromeric histone H3 (CenH3) mediates BubR1 recruitment. *PLoS One* 5(10): e13747.
- Tran, V. et al. (2012). Asymmetric division of *Drosophila* male germline stem cell shows asymmetric histone distribution. *Science* 338(6107): 679-682.
- Treisman, R. and Maniatis, T. (1985). Simian virus 40 enhancer increases number of RNA polymerase II molecules on linked DNA. *Nature* 315(6014): 73-75.
- Trivedi, P. and Stukenberg, P. T. (2015). A Centromere-Signaling Network Underlies the Coordination among Mitotic Events. *Trends Biochem Sci*.
- Tsui, S. et al. (2014). CCCTC-binding factor mediates effects of glucose on beta cell survival. *Cell Prolif* 47(1): 28-37.
- Udvardy, A. et al. (1985). The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J Mol Biol* 185(2): 341-358.
- Val, P. et al. (2003). SF-1 a key player in the development and differentiation of steroidogenic tissues. *Nucl Recept* 1(1): 8.
- Valbuena, A. et al. (2011). Roles of VRK1 as a new player in the control of biological processes required for cell division. *Cell Signal* 23(8): 1267-1272.



## ***Bibliografia***

---

- van Arensbergen, J. et al. (2014). In search of the determinants of enhancer-promoter interaction specificity. *Trends Cell Biol* 24(11): 695-702.
- van Bommel, J. G. et al. (2010). The insulator protein SU(HW) fine-tunes nuclear lamina interactions of the *Drosophila* genome. *PLoS One* 5(11): e15013.
- Van Bortle, K. and Corces, V. G. (2012). tDNA insulators and the emerging role of TFIIC in genome organization. *Transcription* 3(6): 277-284.
- Van Bortle, K. et al. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* 15(6): R82.
- Van Bortle, K. et al. (2012). *Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res* 22(11): 2176-2187.
- van der Horst, A. and Lens, S. M. (2014). Cell division: control of the chromosomal passenger complex in time and space. *Chromosoma* 123(1-2): 25-42.
- Verboon, J. M. et al. (2015). Wash interacts with lamin and affects global nuclear organization. *Curr Biol* 25(6): 804-810.
- Vietri Rudan, M. et al. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep* 10(8): 1297-1309.
- Villa, F. et al. (2009). Crystal structure of the catalytic domain of Haspin, an atypical kinase implicated in chromatin organization. *Proc Natl Acad Sci U S A* 106(48): 20204-20209.
- Villar, D. et al. (2015). Enhancer evolution across 20 mammalian species. *Cell* 160(3): 554-566.
- Villen, J. et al. (2007). Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci U S A* 104(5): 1488-1493.
- Voigt, P. et al. (2013). A double take on bivalent promoters. *Genes Dev* 27(12): 1318-1338.
- Wallace, H. A. and Bosco, G. (2013). Condensins and 3D Organization of the Interphase Nucleus. *Curr Genet Med Rep* 1(4): 219-229.
- Wallace, J. A. and Felsenfeld, G. (2007). We gather together: insulators and genome organization. *Curr Opin Genet Dev* 17(5): 400-407.
- Wang, F. et al. (2012). Haspin inhibitors reveal centromeric functions of Aurora B in chromosome segregation. *J Cell Biol* 199(2): 251-268.
- Wang, F. et al. (2011). A positive feedback loop involving Haspin and Aurora B promotes CPC accumulation at centromeres in mitosis. *Curr Biol* 21(12): 1061-1069.
- Wang, Q. et al. (2015). H3 Thr3 phosphorylation is crucial for meiotic resumption and anaphase onset in oocyte meiosis. *Cell Cycle*: 1-12.

- Wang, Y. et al. (2001). The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell* 105(4): 433-443.
- Wang, Z. et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40(7): 897-903.
- Warren, J. T. et al. (2002). Molecular and biochemical characterization of two P450 enzymes in the ecdysteroidogenic pathway of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 99(17): 11043-11048.
- Warren, J. T. et al. (2004). Phantom encodes the 25-hydroxylase of *Drosophila melanogaster* and *Bombyx mori*: a P450 enzyme critical in ecdysone biosynthesis. *Insect Biochem Mol Biol* 34(9): 991-1010.
- Wendt, K. S. et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451(7180): 796-801.
- Whitfield, W. G. et al. (1995). The 190 kDa centrosome-associated protein of *Drosophila melanogaster* contains four zinc finger motifs and binds to specific sites on polytene chromosomes. *J Cell Sci* 108 ( Pt 11): 3377-3387.
- Wilson-Grady, J. T. et al. (2008). Phosphoproteome analysis of fission yeast. *J Proteome Res* 7(3): 1088-1097.
- Wismar, J. et al. (2000). The mutation without children (*rgl*) causes ecdysteroid deficiency in third-instar larvae of *Drosophila melanogaster*. *Dev Biol* 226(1): 1-17.
- Wood, A. J. et al. (2010). Condensin and cohesin complexity: the expanding repertoire of functions. *Nat Rev Genet* 11(6): 391-404.
- Wood, A. M. et al. (2011). Regulation of chromatin organization and inducible gene expression by a *Drosophila* insulator. *Mol Cell* 44(1): 29-38.
- Woodcock, C. L. and Ghosh, R. P. (2010). Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol* 2(5): a000596.
- Woodman, J. et al. (2015). Phosphorylation of the Scc2 cohesin deposition complex subunit regulates chromosome condensation through cohesin integrity. *Mol Biol Cell* 26(21): 3754-3767.
- Xie, J. et al. (2015). Histone H3 Threonine Phosphorylation Regulates Asymmetric Histone Inheritance in the *Drosophila* Male Germline. *Cell* 163(4): 920-933.
- Xu, Z. et al. (2011). Mapping of INS promoter interactions reveals its role in long-range regulation of SYT8 transcription. *Nat Struct Mol Biol* 18(3): 372-378.
- Yamagishi, Y. et al. (2010). Two histone marks establish the inner centromere and chromosome bi-orientation. *Science* 330(6001): 239-243.

## ***Bibliografía***

---

Yamanaka, N. et al. (2013). Ecdysone control of developmental transitions: lessons from *Drosophila* research. *Annu Rev Entomol* 58: 497-516.

Yanagawa, S. et al. (2002). Casein kinase I phosphorylates the Armadillo protein and induces its degradation in *Drosophila*. *EMBO J* 21(7): 1733-1742.

Yoshiyama, T. et al. (2006). Neverland is an evolutionally conserved Rieske-domain protein that is essential for ecdysone synthesis and insect growth. *Development* 133(13): 2565-2574.

Yuan, W. et al. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem* 286(10): 7983-7989.

Zhang, W. et al. (2003). Genetic and phenotypic analysis of alleles of the *Drosophila* chromosomal JIL-1 kinase reveals a functional requirement at multiple developmental stages. *Genetics* 165(3): 1341-1354.

Zhou, J. and Levine, M. (1999). A novel cis-regulatory element, the PTS, mediates an anti-insulator activity in the *Drosophila* embryo. *Cell* 99(6): 567-575.

Zhou, L. et al. (2014). Polo-like kinase-1 triggers histone phosphorylation by Haspin in mitosis. *EMBO Rep* 15(3): 273-281.

Zink, D. et al. (2004). Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *J Cell Biol* 166(6): 815-825.

Zuin, J. et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci U S A* 111(3): 996-1001.

# Anexo I



## RESEARCH ARTICLE

The insulator protein CTCF regulates *Drosophila* steroidogenesisUjué Fresán<sup>1</sup>, Sergi Cuartero<sup>1,\*</sup>, Michael B. O'Connor<sup>2</sup> and M. Lluïsa Espinàs<sup>1,2,†</sup>

## ABSTRACT

The steroid hormone ecdysone is a central regulator of insect development. In this report we show that CTCF expression in the prothoracic gland is required for full transcriptional activation of the Halloween genes *spookier*, *shadow* and *noppera-bo*, which encode ecdysone biosynthetic enzymes, and for proper timing of ecdysone-responsive gene expression. Loss of CTCF results in delayed and less synchronized larval development that can only be rescued by feeding larvae with both, the steroid hormone 20-hydroxyecdysone and cholesterol. Moreover, CTCF-knockdown in prothoracic gland cells leads to increased lipid accumulation. In conclusion, the insulator protein CTCF is required for Halloween gene expression and cholesterol homeostasis in ecdysone-producing cells controlling steroidogenesis.

**KEY WORDS:** CTCF, Ecdysone, Halloween genes, Developmental timing

## INTRODUCTION

The proper development of multicellular organisms requires accurate timing of specific developmental programs of gene expression. These timed programs involve signaling systems that respond to nutritional and environmental cues to direct coordinated developmental responses throughout the animal, such as transitions in morphology. In insects, these transitions include molting and metamorphosis which occur at regularly defined intervals and depend on pulses of the steroid hormone ecdysone (Yamanaka et al., 2013). During larval and pupal stages of insect development, the prothoracic gland (PG) is the tissue responsible for the synthesis of ecdysone that is secreted and converted into the active steroid molting hormone 20-hydroxyecdysone (20E) in target tissues, where it induces the activation of target genes. The neuropeptide prothoracicotropic hormone (PTTH) is produced by two pairs of lateral neurosecretory cells in the insect brain and is released to signal to the PG (McBrayer et al., 2007). PTTH, through activation of its receptor torso and the mitogen-activated protein kinase (MAPK) cascade (Rewitz et al., 2009), leads to increased transcription of ecdysone biosynthetic enzymes encoded by the Halloween gene family (Gilbert, 2004). Transcriptional regulation of these genes is still poorly understood, although some transcription factors and chromatin modifiers have recently been identified for some of the biosynthetic genes (Danielsen et al.,

2014; Deng and Kerppola, 2013; Moeller et al., 2013; Pankotai et al., 2010; Parvy et al., 2005). Halloween genes trigger the production of ecdysone from dietary cholesterol or phytosterols, since arthropods are unable to synthesize cholesterol. In recent years, evidence has accumulated that other factors in addition to PTTH, especially nutritional signals, act on the PG to control ecdysteroidogenesis (Layalle et al., 2008).

The insulator protein CTCF is a highly conserved zinc finger protein that has been shown to play an essential role in regulation of chromatin organization and gene expression during development both, in *Drosophila* and mammals (Herold et al., 2012). Recent data supports a role for insulator proteins in the organization of developmentally regulated intra- and interchromosomal contacts that are responsible for the diverse functions of CTCF in gene regulation, including context-dependent transcriptional activation, repression, insulation and imprinting (Phillips and Corces, 2009).

In this paper we show that CTCF is required for Halloween gene expression and cholesterol homeostasis in the prothoracic gland controlling steroidogenesis in *Drosophila*.

## RESULTS AND DISCUSSION

## CTCF controls the timing of larval development

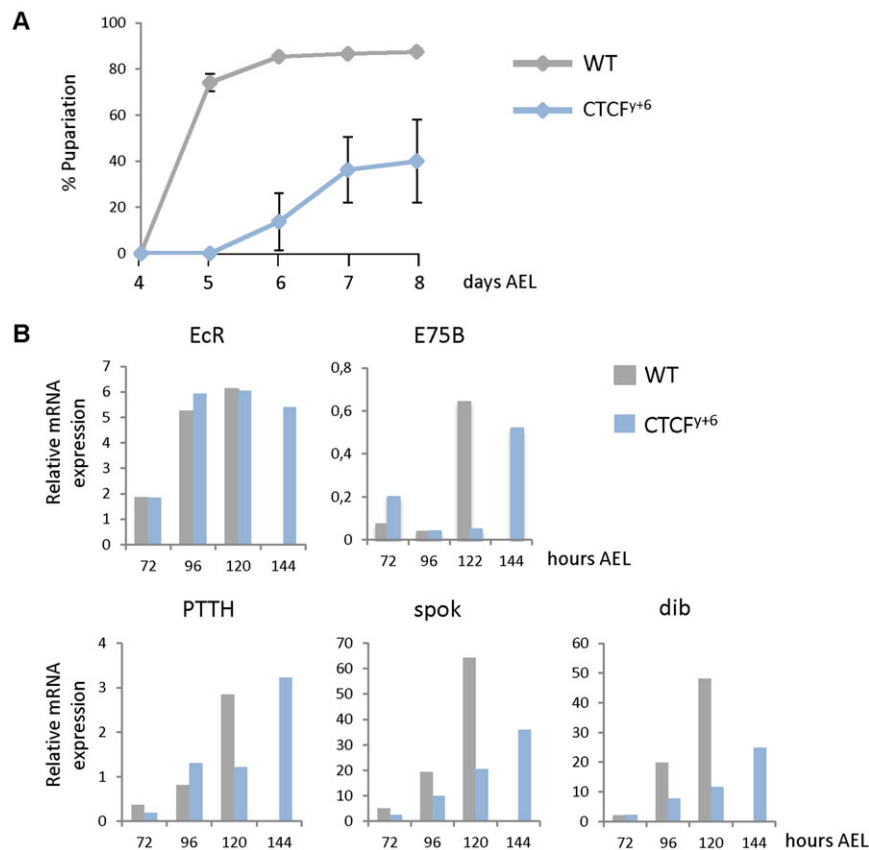
The analysis of *Drosophila* developmental timing in *CTCF<sup>+/+</sup>* null mutant background (Gerasimova et al., 2007) revealed that larval development was prolonged by two days on average in those mutant animals that progressed to the larval/pupal transition (Fig. 1A). We confirmed that the developmental delay in these animals was due to the loss of CTCF by performing rescue experiments using the UAS/Gal4 system. It has been reported that expression of *UAS-CTCF* with ubiquitous drivers lead to embryonic lethality while a single heat-shock during larval development of animals containing a *hsp70-Gal4* transgene (*hs>*) was sufficient to phenotypically rescue the majority of CTCF mutant flies (Mohan et al., 2007). Our analyses showed that a short pulse of CTCF significantly rescued the developmental delay in *CTCF<sup>+/+</sup>* mutants (supplementary material Fig. S1). Indeed, although 90% of *CTCF<sup>+/+</sup>* larvae died before pupariation after a 20 min heat-shock, some larvae did reach the pupal stage but pupariation was delayed by more than 3 days. CTCF expression was able to partially rescue both, lethality (around 75% of mutant larvae pupariate, data not shown) and, importantly, developmental delay (pupariation was clearly advanced by CTCF expression in the *CTCF<sup>+/+</sup>* mutant background, supplementary material Fig. S1).

The delayed pupariation of *CTCF<sup>+/+</sup>* mutants suggested that the ecdysone signaling pathway might be affected in these animals. Since the insulator protein CTCF plays an essential role in regulation of gene expression, we analyzed the transcript levels of components acting at different steps in ecdysone signaling pathway. It has been previously reported that the ecdysone-induced protein 75B locus contains a CTCF site involved in recruitment of the insulator protein CP190 in response to the hormone and proper activation of gene expression (Wood et al., 2011). Thus, we analyzed the expression of the ecdysone-responsive gene *E75B* during larval development. We also

<sup>1</sup>Institute of Molecular Biology of Barcelona, IBMB-CSIC, and Institute for Research in Biomedicine IRB, Barcelona 08028, Spain. <sup>2</sup>Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, MN 55455, USA. \*Present address: MRC Clinical Sciences Centre, Imperial College London, London W12 0NN, UK

†Author for correspondence (mlebm@ibmb.csic.es)

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.



**Fig. 1. Loss of CTCF produces developmental delay.**

(A) The percentages of larvae of the indicated genotypes that underwent pupariation were plotted relative to the time in days after egg laying. Error bars show s.d. of three independent experiments. (B) Representative transcriptional profiles along time in hours after egg laying of the indicated genes in wild type and CTCF<sup>Y+6</sup> mutant larvae.

analyzed the transcript levels of the Ecdysone receptor (EcR), which is activated by hormone to regulate expression of ecdysone-responsive genes, PTTH, which is the peptide hormone that initiates steroid production, and the Halloween genes *spookier* (*spok*) (Ono et al., 2006) and *disembodied* (*dib*) (Chavez et al., 2000), two ecdysone biosynthetic enzymes expressed in the PG. Our analyses showed no changes in *EcR* transcriptional levels in CTCF mutant animals and only a delay in the induction of the expression of *E75B* and *PTTH*, each of which eventually showed transcript levels similar to controls (Fig. 1B). However, we observed a reduction in the transcriptional induction of two genes involved in ecdysone biosynthesis, *spok* and *dib*. These results suggest that CTCF might regulate induction of certain Halloween genes necessary for ecdysone synthesis.

#### Loss of CTCF in ecdysone-producing cells impairs Halloween gene transcriptional activation

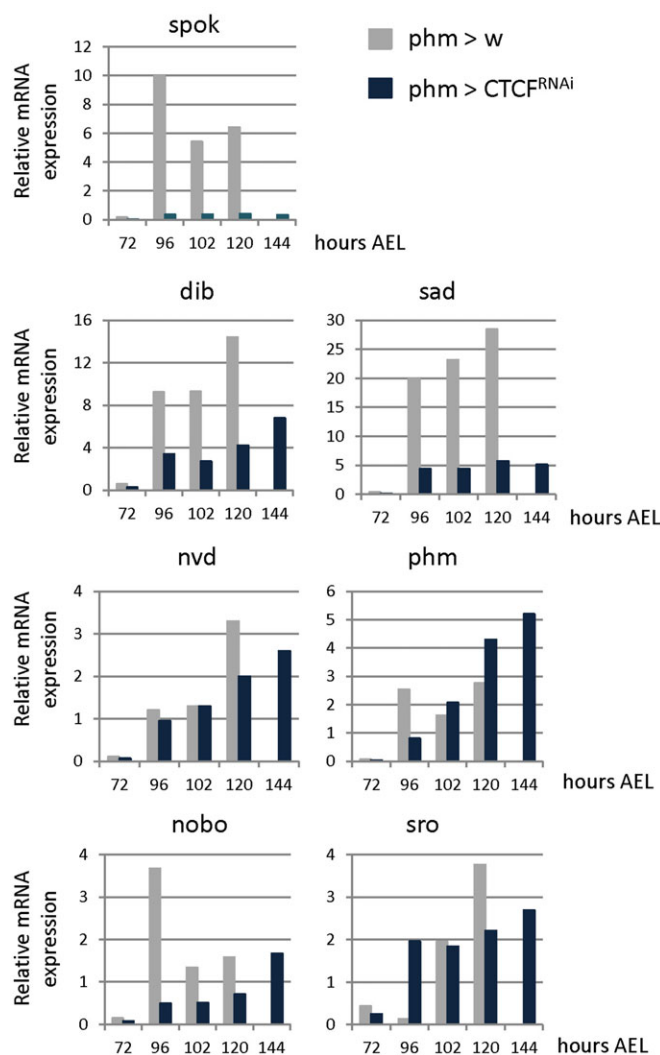
In order to analyze whether CTCF is required autonomously in the ecdysone-producing cells themselves to regulate Halloween gene expression we knocked down CTCF specifically in the PG using the UAS/Gal4 system. A *UAS-CTCF-RNAi* construct combined with *UAS-dicer2* was expressed in the prothoracic cells using a PG-specific *phantom-Gal4* (*phm*>) driver. We then measured the transcript levels of the Halloween genes *spok*, *dib*, *shadow* (*sad*) (Warren et al., 2002), *neverland* (*nvd*) (Yoshiyama et al., 2006), *phantom* (*phm*) (Niwa et al., 2004; Warren et al., 2004), *noppera-bo* (*nobo*) (Enya et al., 2014) and *shroud* (*sro*) (Niwa et al., 2010). Control larvae that only express *phantom-Gal4* in a wild type background exhibit a clear expression increase of these genes during larval development (Fig. 2). The expression of these Halloween genes is initially quite low during the early third-instar stage and then rises, as previously reported (McBrayer et al., 2007). In

*phm*>*CTCF<sup>RNAi</sup>* background the transcript levels of *dib*, *sad* and *nobo* were 2-5 fold lower in *phm*>*CTCF<sup>RNAi</sup>* larvae at the time points of maximal induction (Fig. 2) and we observed almost no induction of *spok*, which is one of the genes of the so-called Black-Box that has been suggested to be a rate limiting step in ecdysone synthesis (Warren et al., 2009). On the other hand, no clear differences were observed for the Halloween genes *nvd*, *phm* and *sro* indicating that CTCF is required specifically in ecdysone-producing cells for activation of several, but not all of the ecdysone biosynthetic enzymes.

We investigated the possibility that loss of CTCF in the PG might indirectly affect ecdysone production by modifying PTTH signaling. First we analyzed whether the gene is expressed normally and found that *PTTH* transcripts are only found in the PG neurons in each brain hemisphere as in controls and the transcript levels were not decreased. In fact they were actually increased, during the prolonged larval stage (supplementary material Fig. S2A and B). We also analyzed the expression of the receptor *torso* and found similar transcript levels in *phm*>*CTCF<sup>RNAi</sup>* knockdown as in control larvae (supplementary material Fig. S2C). Therefore, our data indicate that loss of CTCF in the PG does not likely interfere with PTTH signaling to the PG.

#### CTCF expression in the PG is required for proper timing of ecdysone-signaling gene expression

Reduced levels of CTCF in the PG led to partial lethality at different larval stages, but many knockdown animals were able to pupariate although most died before eclosion. We analyzed developmental timing in *phm*>*CTCF<sup>RNAi</sup>* knockdown larvae and we found that larval development was prolonged by two days on average and was less synchronous than control animals (Fig. 3A). We also measured



**Fig. 2. CTCF knockdown in the prothoracic gland regulates expression of the Halloween genes.** Representative transcriptional profiles versus time in hours after egg laying of the Halloween genes *spok*, *dib*, *sad*, *nvd*, *phm*, *nobo* and *sro* in control (*phm*>*w*) and CTCF RNAi (*phm*>*CTCF<sup>RNAi</sup>*) larvae.

the average time to ecdysis from the second to the third instar stage and found that it was increased by around 8 h for CTCF knockdown animals (supplementary material Fig. S3).

Since larval development was prolonged by two-three days in *phm*>*CTCF<sup>RNAi</sup>* knockdown animals, we analyzed Halloween gene expression at 168 and 192 h AEL. We found that transcript levels of *dib*, *nvd*, *phm* and *sro* increased as in control animals during the prolonged larval development (Fig. 3B). In contrast, even though at the end of larval development there was some increase in *spok*, *sad* and *nobo* transcription, transcript levels of these genes were always significantly lower in *phm*>*CTCF<sup>RNAi</sup>* compared to control animals (Fig. 3B). Moreover, difference for *spok* and *nobo* gene expression between control and CTCF knockdown animals is even higher if one considers that in control larva maximum expression is seen at 96 h and not 120 h of larval development (Fig. 2).

Interestingly, despite low transcriptional levels of some ecdysone biosynthetic genes in *phm*>*CTCF<sup>RNAi</sup>* animals during larval development, some animals are able to eventually mount a sufficient hormone response to initiate metamorphosis. To analyze whether there is a perturbation of ecdysone accumulation in these animals, we measured the transcript levels of the ecdysone-responsive

gene *E74B* as a proxy for the ecdysone titer during the larval stage. In control larvae transcript levels of *E74B* showed a peak at 96 h AEL while in CTCF knockdown larvae a similar peak was reached but two days later (Fig. 3C), consistent with the length of the developmental delay exhibited by these mutant animals. These results are similar to that seen in animals lacking PTTH which mount a proper, but asynchronous, transcriptional response to ecdysone (McBrayer et al., 2007). Taken together our results indicate that the production of ecdysone, not the response to the hormone, is the source of delay in CTCF knockdown animals.

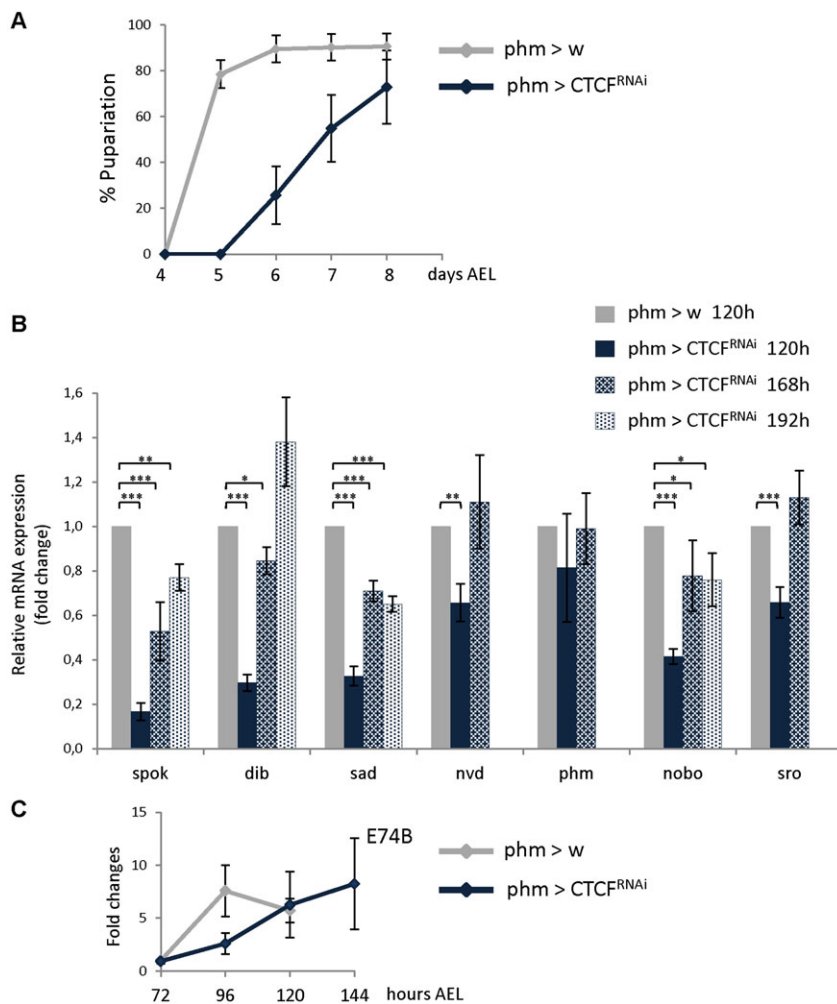
We also analyzed the participation of the insulator proteins CP190 and Ibf1/2 whose binding sites were previously shown to significantly overlap with CTCF-binding sites (Cuartero et al., 2014). We found that developmental timing was normal in both *phm*>*CP190<sup>RNAi</sup>* and *phm*>*Ibf1/2<sup>RNAi</sup>* larvae (data not shown), suggesting that the CP190-Ibf1/2 insulator complex does not participate in CTCF-mediated regulation of Drosophila steroidogenesis.

### Developmental timing depends on CTCF-mediated cholesterol homeostasis and ecdysone synthesis

Though *phm*>*CTCF<sup>RNAi</sup>* knockdown larvae showed low transcriptional levels of the Halloween genes *spok*, *sad* and *nobo*, they are able to induce ecdysone-responsive genes such as *E74B* and most larvae still progressed to the larval/pupal transition. However, pupariation was delayed in these animals, suggesting that CTCF is needed for proper timing of high level ecdysone production. To determine if low levels of ecdysteroids are responsible for the developmental delay, we fed either 20-hydroxyecdysone (20E), which is the active form of the hormone, or cholesterol (C), which is the precursor for ecdysone biosynthesis in the PG, to *phm*>*CTCF<sup>RNAi</sup>*. Although some of the 20E-fed CTCF-knockdown larvae started to pupariate at the same time as control animals, individuals of the *phm*>*CTCF<sup>RNAi</sup>* population pupariate over a longer period of time (Fig. 4A). Since it has been recently reported that Ecdysone (E) and 20E have different effects on developmental timing of wild type larvae (Ono, 2014), we also fed E to *phm*>*CTCF<sup>RNAi</sup>* larvae but we did not observe any differences with respect to 20E-fed larvae (data not shown). On the other hand, when larvae were cholesterol-fed puparium formation and adult eclosion were normal although delayed by an average of one day. Indeed, we were only able to efficiently rescue developmental delay and adult eclosion by feeding larvae with both, 20E and C, suggesting that CTCF is needed not only to synthesize ecdysone from cholesterol but also to mediate cholesterol homeostasis in PG cells.

It has been reported that at the onset of pupariation there is an increase of lipid droplets in PG cells that can be seen by Oil Red O staining of precisely staged 120 h AEL wandering larvae (Talamillo et al., 2013). At 120 h AEL, *phm*>*CTCF<sup>RNAi</sup>* PGs had a reduced content of lipid droplets in comparison to controls (Fig. 4B, compare middle and left panels), likely due to developmental delay. However, there is also an increase of lipid droplets at the end of larval development in *phm*>*CTCF<sup>RNAi</sup>* PG cells which tend to be slightly higher than in control animals (Fig. 4B, compare right and left panels and supplementary material Fig. S4). Since the subcellular lipid accumulation phenotype of *phm*>*CTCF<sup>RNAi</sup>* PG cells is similar to that of Niemann-Pick type C (*npc*) mutants (Huang et al., 2005), we analyzed *dnpc1a* transcriptional levels but we found no changes between control and *phm*>*CTCF<sup>RNAi</sup>* larvae (data not shown). Increased lipid accumulation in the fat body in EcR knockdown larvae has been reported (Kamoshida et al., 2012), but our results show that *EcR* transcriptional levels do not change in





**Fig. 3. The increase in ecdysone signaling is delayed in CTCF knockdown larvae.** (A) The percentages of larvae of the control (phm>w) and CTCF RNAi (phm>CTCF<sup>RNAi</sup>) genotypes that underwent pupariation were plotted relative to the time in days after egg laying. Error bars show s.d. of three independent experiments. (B) Transcriptional levels of Halloween genes in CTCF RNAi at the end of larval development as fold changes relative to control (phm>w 120 h). Error bars show s.e.m. of at least three independent experiments. \*\*\* $P<0.001$ , \*\* $P<0.01$ , \* $P<0.1$  with Student's *t*-test. (C) Temporal profile of ecdysone signaling gene *E74B* transcriptional induction in control and CTCF RNAi larvae. Error bars show s.e.m. of three independent experiments.

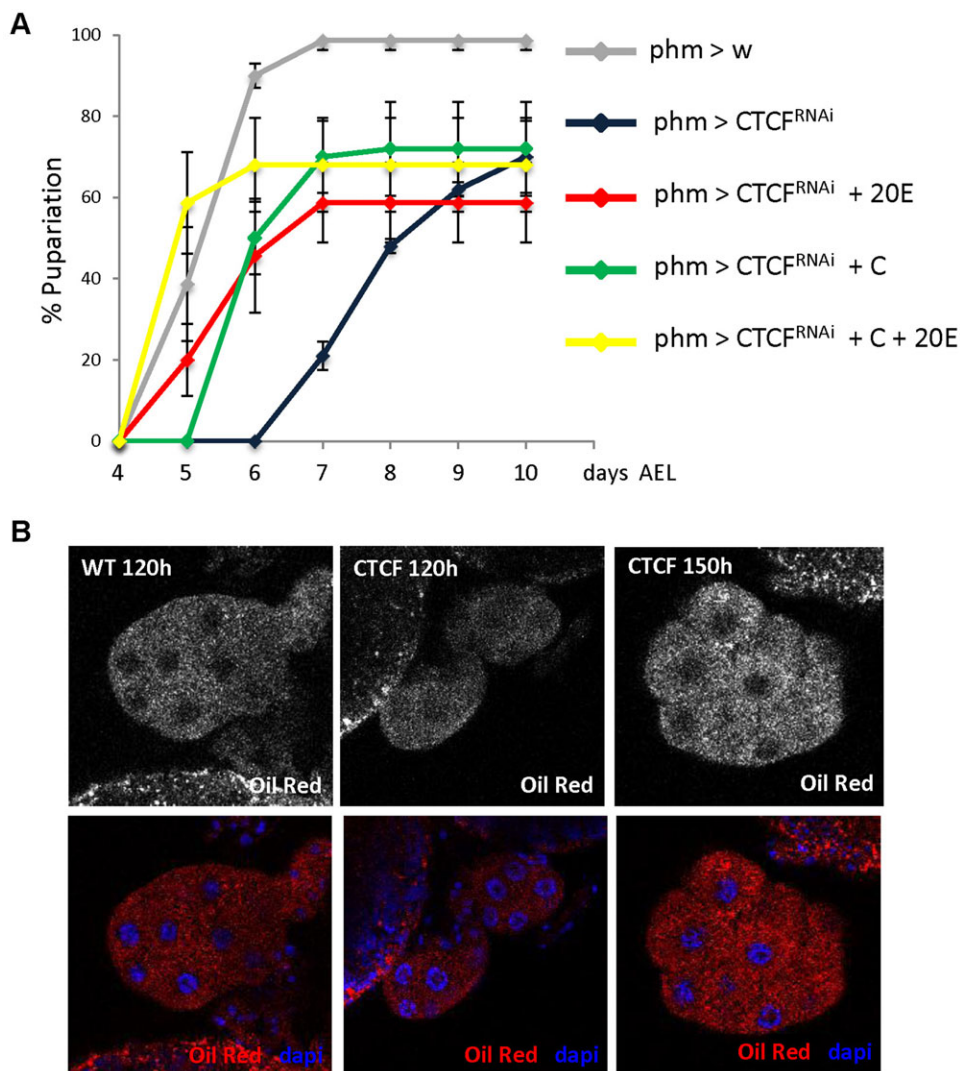
CTCF mutant animals (Fig. 1B). On the other hand, as shown above, CTCF expression in the prothoracic gland is required for transcriptional activation of the Halloween gene *nobo* (Fig. 2), which has been shown to be necessary for normal accumulation of cholesterol in PG cells (Enya et al., 2014). Thus, the abnormal accumulation of lipid droplets in CTCF knockdown animals are in agreement with CTCF-mediated regulation of *nobo* gene expression. These data along with results reported above showing partial rescue in cholesterol-fed CTCF-knockdown larvae (Fig. 4A) suggest a role of CTCF in sterol homeostasis in the PG.

In recent years evidence has accumulated that PTH is a major, but not the only, developmental signal that triggers ecdysteroidogenesis; multiple factors, in particular nutritional signals reflecting the general metabolic status of the larvae, act directly on the PG to control ecdysteroidogenesis. Insulin and Target of Rapamycin (TOR) pathways have been shown to play a direct role in regulating ecdysone synthesis in the PG (Caldwell et al., 2005; Colombani et al., 2005; Layalle et al., 2008; Mirth et al., 2005). Insulin signaling has been studied extensively in *Drosophila* as a regulator of tissue growth and involves activation of several downstream factors including PI3K and AKT. In contrast, PTH signals through activation of the receptor Torso and the Erk pathway (Rewitz et al., 2009). Interestingly, it has been reported that CTCF plays a role in insulin-induced cell proliferation and acts downstream of insulin-induced activation of both Erk and Akt (Gao et al., 2007; Tsui et al., 2014). In addition to insulin/TOR, recent

work has shown that TGF $\beta$ /activin signaling is also essential for ecdysone synthesis and developmental transitions (Gibbens et al., 2011). It has been suggested that *dSmad2*, the major downstream mediator of TGF $\beta$  signals, might participate directly in regulating Halloween gene transcription. The observation that TGF $\beta$  promotes complexes between Smad proteins and CTCF on the H19 imprinting control region (Bergstrom et al., 2010), raises the possibility that these complexes could also regulate Halloween gene transcription. We searched for CTCF-binding sites in the promoter regions of Halloween genes and we found motifs at less than -1 kb from TSSs in all of them except *nvd* (supplementary material Fig. S5). Nevertheless, it remains to be determined whether CTCF-mediated induction of these genes is mediated through these sites and whether complexes with other transcription factors might also be involved.

## Conclusions

Here we show that the insulator protein CTCF plays important roles in *Drosophila* steroidogenesis at two different levels: on one hand, CTCF controls cholesterol homeostasis in the PG and, on the other hand, CTCF is required for activation of the enzymes responsible for the synthesis of the hormone. Whether these two events are coupled needs further analyses but our data suggest that this is unlikely since to efficiently rescue of developmental delay in CTCF-knockdown larvae required both, ecdysone and cholesterol. Overall, our results suggest that the insulator protein CTCF might be part of a molecular mechanism that couple nutritional input and timing of juvenile maturation.



**Fig. 4. Both, 20-hydroxyecdysone (20E) and cholesterol (C) are required to rescue CTCF knockdown phenotypes.** (A) Control (*phm>w*) and CTCF RNAi (*phm>CTCF<sup>RNAi</sup>*) larvae were fed food containing EtOH, 20E or/and C and the percentages of larvae of the indicated genotypes that underwent pupariation were plotted relative to the time in days after egg laying. Error bars show s.d. of three independent experiments. (B) Lipid accumulation in control and CTCF RNAi ecdysone-producing cells. Single plane confocal micrographs showing lipid droplets in PG cells stained with Oil Red O (either in black and white, upper panels, or in red, lower panels) of 120 h AEL clear-gut *phm>w* control larvae (left panels), 120 h AEL blue-gut *phm>CTCF<sup>RNAi</sup>* larvae (middle panels) and 150 h AEL clear-gut *phm>CTCF<sup>RNAi</sup>* larvae (right panels). DNA was labeled with DAPI (blue).

## MATERIALS AND METHODS

### Fly stocks

All flies were raised at 25°C on standard medium. The following transgenic and mutant flies were used: *CTCF<sup>γ+6</sup>* (Gerasimova et al., 2007), *phm-Gal4* (Rewitz et al., 2009), *hsp70-Gal4* (Bloomington Drosophila Stock Center #2077), *UAS-dCTCF* (Mohan et al., 2007) *UAS-CTCF-RNAi* (Vienna Drosophila RNAi Center-VDR #30713), *UAS-CP190-RNAi* (VDR #35078), *UAS-Ibf1-RNAi* (VDR #35426), *UAS-Ibf2-RNAi* (VDR #42121) and *UAS-dicer2* (VDR #60008).

### Staging larvae, developmental timing analyses and feeding experiments

Females were allowed to lay eggs on apple juice agar plates for 2 h. Synchronized new hatched first-instar larvae were placed on standard medium supplemented with yeast at 25°C and allowed to develop to the desired time points. For developmental timing analyses, 30 larvae were placed per vial and pupariation timing was scored periodically. A single heat shock (37°C) for 20 min at 24 h of development was performed in rescue experiments. Feeding experiments were performed by adding 20-Hydroxyecdysone (Santa Cruz) at 0.33 mg/ml in ethanol and/or Cholesterol (Sigma) at 0.35 mg/ml. Control larvae were fed with ethanol.

### Quantitative RT-PCR

Total RNAs were extracted from ring gland+brain of precisely staged larvae with TRIzol Reagent (Life Technologies) and purified with RNeasy Mini Kit (Qiagen). cDNAs were prepared from 0.5 µg of RNA using the

Transcriptor First Strand cDNA Synthesis Kit (Roche). qRT-PCR was performed on Roche LightCycler 480 System (Roche) using LightCycler 480 SYBER Green I Master (Roche). Three independent biological replicates for each genotype at each time point were performed. Transcriptional levels were normalized to ribosomal protein Rpl23. Primers used are listed in supplementary material Table S1.

### Oil Red O staining

Ring gland+brain were fixed in 4% paraformaldehyde for 20 min, washed twice in PBS and stained with Oil Red O (Sigma) solution at 0.06% in isopropanol for 30 min. Samples were washed and stained with DAPI before mounting in Vectashield (Vector Laboratories). Single plane images were taken with a confocal Leica TCS SP2-AOBS microscope.

### Acknowledgements

We thank David Martin and Josep Casacuberta for helpful discussions and comments on the manuscript. We also thank members of Michael O'Connor laboratory for materials and technical support. We are thankful to Victor Corces, Rainer Renkawitz, Bloomington Drosophila Stock Center and Vienna Drosophila RNAi Center for providing fly stocks. U.F. and S.C. acknowledge receipt of doctoral fellowships from CSIC and MEC respectively.

### Competing interests

The authors declare no competing or financial interests.

### Author contributions

M.L.E. designed the study. U.F., S.C. and M.L.E. performed the experiments. U.F., S.C., M.B.O. and M.L.E. analyzed the data. M.L.E. wrote the manuscript.

## Funding

This work was supported by grants from the Spanish Ministerio de Ciencia e Innovación (BFU2009-10983 and BFU2013-48712-P). M.B.O. was supported by National Institutes of Health grant R01 GM093301.

## Supplementary material

Supplementary material available online at  
<http://bio.biologists.org/lookup/suppl/doi:10.1242/bio.012344/-/DC1>

## References

- Bergstrom, R., Savary, K., Moren, A., Guibert, S., Heldin, C.-H., Ohlsson, R. and Moustakas, A. (2010). Transforming growth factor beta promotes complexes between Smad proteins and the CCCTC-binding factor on the H19 imprinting control region chromatin. *J. Biol. Chem.* **285**, 19727-19737.
- Caldwell, P. E., Walkiewicz, M. and Stern, M. (2005). Ras activity in the Drosophila prothoracic gland regulates body size and developmental rate via ecdysone release. *Curr. Biol.* **15**, 1785-1795.
- Chavez, V. M., Marques, G., Delbecq, J. P., Kobayashi, K., Hollingsworth, M., Burr, J., Natze, J. E. and O'Connor, M. B. (2000). The Drosophila disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels. *Development* **127**, 4115-4126.
- Colombani, J., Bianchini, L., Layalle, S., Pondeville, E., Dauphin-Villemant, C., Antoniewski, C., Carré, C., Noselli, S. and Léopold, P. (2005). Antagonistic actions of ecdysone and insulins determine final size in Drosophila. *Science* **310**, 667-670.
- Cuartero, S., Fresan, U., Reina, O., Planet, E. and Espinas, M. L. (2014). Ibf1 and Ibf2 are novel CP190-interacting proteins required for insulator function. *EMBO J.* **33**, 637-647.
- Danielsen, E. T., Moeller, M. E., Dorry, E., Komura-Kawa, T., Fujimoto, Y., Troelsen, J. T., Herder, R., O'Connor, M. B., Niwa, R. and Rewitz, K. F. (2014). Transcriptional control of steroid biosynthesis genes in the Drosophila prothoracic gland by ventral veins lacking and knirps. *PLoS Genet.* **10**, e1004343.
- Deng, H. and Kerppola, T. K. (2013). Regulation of Drosophila metamorphosis by xenobiotic response regulators. *PLoS Genet.* **9**, e1003263.
- Enya, S., Ameku, T., Igarashi, F., Iga, M., Kataoka, H., Shinoda, T. and Niwa, R. (2014). A Halloween gene noppera-bo encodes a glutathione S-transferase essential for ecdysteroid biosynthesis via regulating the behaviour of cholesterol in Drosophila. *Sci. Rep.* **4**, 6586.
- Gao, J., Li, T. and Lu, L. (2007). Functional role of CCCTC binding factor in insulin-stimulated cell proliferation. *Cell Prolif.* **40**, 795-808.
- Gerasimova, T. I., Lei, E. P., Bushey, A. M. and Corces, V. G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. *Mol. Cell* **28**, 761-772.
- Gibbens, Y. Y., Warren, J. T., Gilbert, L. I. and O'Connor, M. B. (2011). Neuroendocrine regulation of Drosophila metamorphosis requires TGF $\beta$ /Activin signaling. *Development* **138**, 2693-2703.
- Gilbert, L. I. (2004). Halloween genes encode P450 enzymes that mediate steroid hormone biosynthesis in Drosophila melanogaster. *Mol. Cell. Endocrinol.* **215**, 1-10.
- Herold, M., Bartkuhn, M. and Renkawitz, R. (2012). CTCF: insights into insulator function during development. *Development* **139**, 1045-1057.
- Huang, X., Suyama, K., Buchanan, J., Zhu, A. J. and Scott, M. P. (2005). A Drosophila model of the Niemann-Pick type C lysosome storage disease: dnpc1a is required for molting and sterol homeostasis. *Development* **132**, 5115-5124.
- Kamoshida, Y., Fujiyama-Nakamura, S., Kimura, S., Suzuki, E., Lim, J., Shiozaki-Sato, Y., Kato, S. and Takeyama, K.-i. (2012). Ecdysone receptor (EcR) suppresses lipid accumulation in the Drosophila fat body via transcription control. *Biochem. Biophys. Res. Commun.* **421**, 203-207.
- Layalle, S., Arquier, N. and Léopold, P. (2008). The TOR pathway couples nutrition and developmental timing in Drosophila. *Dev. Cell* **15**, 568-577.
- McBrayer, Z., Ono, H., Shimell, M., Parvy, J.-P., Beckstead, R. B., Warren, J. T., Thummel, C. S., Dauphin-Villemant, C., Gilbert, L. I. and O'Connor, M. B. (2007). Prothoracicotrophic hormone regulates developmental timing and body size in Drosophila. *Dev. Cell* **13**, 857-871.
- Mirth, C., Truman, J. W. and Riddiford, L. M. (2005). The role of the prothoracic gland in determining critical weight for metamorphosis in Drosophila melanogaster. *Curr. Biol.* **15**, 1796-1807.
- Moeller, M. E., Danielsen, E. T., Herder, R., O'Connor, M. B. and Rewitz, K. F. (2013). Dynamic feedback circuits function as a switch for shaping a maturation-inducing steroid pulse in Drosophila. *Development* **140**, 4730-4739.
- Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heini, N., Bardenhagen, I., Leers, J., White, R. A. H., Renkawitz-Pohl, R., Saumweber, H. et al. (2007). The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J.* **26**, 4203-4214.
- Niwa, R., Matsuda, T., Yoshiyama, T., Namiki, T., Mita, K., Fujimoto, Y. and Kataoka, H. (2004). CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid biosynthesis in the prothoracic glands of Bombyx and Drosophila. *J. Biol. Chem.* **279**, 35942-35949.
- Niwa, R., Namiki, T., Ito, K., Shimada-Niwa, Y., Kiuchi, M., Kawaoka, S., Kayukawa, T., Banno, Y., Fujimoto, Y., Shigenobu, S. et al. (2010). Non-molting glossy/shroud encodes a short-chain dehydrogenase/reductase that functions in the 'Black Box' of the ecdysteroid biosynthesis pathway. *Development* **137**, 1991-1999.
- Ono, H. (2014). Ecdysone differentially regulates metamorphic timing relative to 20-hydroxyecdysone by antagonizing juvenile hormone in Drosophila melanogaster. *Dev. Biol.* **391**, 32-42.
- Ono, H., Rewitz, K. F., Shinoda, T., Itoyama, K., Petryk, A., Rybczynski, R., Jarcho, M., Warren, J. T., Marqués, G., Shimell, M. J. et al. (2006). Spook and Spookier code for stage-specific components of the ecdysone biosynthetic pathway in Diptera. *Dev. Biol.* **298**, 555-570.
- Pankotai, T., Popescu, C., Martin, D., Grau, B., Zsindely, N., Bodai, L., Tora, L., Ferrus, A. and Boros, I. (2010). Genes of the ecdysone biosynthesis pathway are regulated by the dATAC histone acetyltransferase complex in Drosophila. *Mol. Cell. Biol.* **30**, 4254-4266.
- Parvy, J.-P., Blais, C., Bernard, F., Warren, J. T., Petryk, A., Gilbert, L. I., O'Connor, M. B. and Dauphin-Villemant, C. (2005). A role for betaFTZ-F1 in regulating ecdysteroid titers during post-embryonic development in Drosophila melanogaster. *Dev. Biol.* **282**, 84-94.
- Phillips, J. E. and Corces, V. G. (2009). CTCF: master weaver of the genome. *Cell* **137**, 1194-1211.
- Rewitz, K. F., Yamanaka, N., Gilbert, L. I. and O'Connor, M. B. (2009). The insect neuropeptide PTTH activates receptor tyrosine kinase torso to initiate metamorphosis. *Science* **326**, 1403-1405.
- Talamillo, A., Herbozo, L., Pirone, L., Pérez, C., González, M., Sánchez, J., Mayor, U., Lopitz-Otsoa, F., Rodríguez, M. S., Sutherland, J. D. et al. (2013). Scavenger receptors mediate the role of SUMO and Ftz-f1 in Drosophila steroidogenesis. *PLoS Genet.* **9**, e1003473.
- Tsui, S., Dai, W. and Lu, L. (2014). CCCTC-binding factor mediates effects of glucose on beta cell survival. *Cell Prolif.* **47**, 28-37.
- Warren, J. T., Petryk, A., Marques, G., Jarcho, M., Parvy, J.-P., Dauphin-Villemant, C., O'Connor, M. B. and Gilbert, L. I. (2002). Molecular and biochemical characterization of two P450 enzymes in the ecdysteroidogenic pathway of Drosophila melanogaster. *Proc. Natl. Acad. Sci. USA* **99**, 11043-11048.
- Warren, J. T., Petryk, A., Marqués, G., Parvy, J.-P., Shinoda, T., Itoyama, K., Kobayashi, J., Jarcho, M., Li, Y., O'Connor, M. B. et al. (2004). Phantom encodes the 25-hydroxylase of Drosophila melanogaster and Bombyx mori: a P450 enzyme critical in ecdysone biosynthesis. *Insect Biochem. Mol. Biol.* **34**, 991-1010.
- Warren, J. T., O'Connor, M. B. and Gilbert, L. I. (2009). Studies on the Black Box: incorporation of 3-oxo-7-dehydrocholesterol into ecdysteroids by Drosophila melanogaster and Manduca sexta. *Insect Biochem. Mol. Biol.* **39**, 677-687.
- Wood, A. M., Van Bortle, K., Ramos, E., Takenaka, N., Rohrbach, M., Jones, B. C., Jones, K. C. and Corces, V. G. (2011). Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. *Mol. Cell* **44**, 29-38.
- Yamanaka, N., Rewitz, K. F. and O'Connor, M. B. (2013). Ecdysone control of developmental transitions: lessons from Drosophila research. *Annu. Rev. Entomol.* **58**, 497-516.
- Yoshiyama, T., Namiki, T., Mita, K., Kataoka, H. and Niwa, R. (2006). Neverland is an evolutionally conserved Rieske-domain protein that is essential for ecdysone synthesis and insect growth. *Development* **133**, 2565-2574.

# Anexo II





# Ibf1 and Ibf2 are novel CP190-interacting proteins required for insulator function

Sergi Cuartero<sup>1</sup>, Ujué Fresán<sup>1</sup>, Oscar Reina<sup>2</sup>, Evarist Planet<sup>2</sup> & M Lluïsa Espinàs<sup>1,\*</sup>

## Abstract

Insulators are DNA-protein complexes that play a central role in chromatin organization and regulation of gene expression. In *Drosophila* different proteins, dCTCF, Su(Hw), and BEAF bind to specific subsets of insulators most of them having in common CP190. It has been shown that there are a number of CP190-binding sites that are not shared with any other known insulator protein, suggesting that other proteins could cooperate with CP190 to regulate insulator activity. Here we report on the identification of two previously uncharacterized proteins as CP190-interacting proteins, that we have named Ibf1 and Ibf2. These proteins localize at insulator bodies and associate with chromatin at CP190-binding sites throughout the genome. We also show that Ibf1 and Ibf2 are DNA-binding proteins that form hetero-oligomers that mediate CP190 binding to chromatin. Moreover, Ibf1 and Ibf2 are necessary for insulator activity in enhancer-blocking assays and Ibf2 null mutation cause a homeotic phenotype. Taken together our data reveal a novel pathway of CP190 recruitment to chromatin that is required for insulator activity.

**Keywords** CP190; CTCF; Fab-8; insulator

**Subject Categories** Chromatin, Epigenetics, Genomics & Functional Genomics; Transcription

**DOI** 10.1002/embj.201386001 | Received 17 June 2013 | Revised 14 November 2013 | Accepted 18 December 2013 | Published online 6 February 2014

**EMBO Journal (2014) 33, 637–647**

## Introduction

Insulators are DNA-protein complexes that are thought to play a role in chromatin organization and the regulation of gene expression by mediating intra- and interchromosomal interactions (Maeda & Karch, 2007; Wallace & Felsenfeld, 2007; Van Bortle & Corces, 2013). Insulators have been characterized in all organisms, from yeast to humans (West *et al*, 2002; Valenzuela & Kamakaka, 2006). In *Saccharomyces cerevisiae* and *S. pombe* insulators are primarily promoters of highly active RNA pol II and RNA pol III transcribed genes or TFIIIC-bound sequences that block the spread of silencing (Oki & Kamakaka, 2005; Noma *et al*, 2006). In metazoans, although

there are some examples of gene promoters functioning as insulators (Lunyak *et al*, 2007; Chopra *et al*, 2009) and it has recently been shown that tRNA genes have insulator activity (Ebersole *et al*, 2011; Raab *et al*, 2012), most of the insulators are autonomous DNA-protein complexes. Vertebrates possess only one known insulator protein, CTCF, while in *Drosophila* different proteins, dCTCF, Su(Hw), BEAF and GAGA, bind to the different insulators characterized. In addition, CP190 protein seems to be common to all of them (Gurudatta & Corces, 2009; Ahanger *et al*, 2013). Several studies indicate that insulator elements and the proteins they bind play a fundamental role in development and disease. In *Drosophila*, the bithorax and antennapedia complexes contain several transcriptional regulatory sequences that orchestrate the complex spatio-temporal expression of the homeotic genes present in these regions. Several studies have shown that the proper interplay between these regulatory sequences requires the function of different insulators (Karch *et al*, 1994; Barges *et al*, 2000; Belozarov *et al*, 2003; Perez-Lluch *et al*, 2008). Recently, CTCF occupancy was mapped in different cell types and showed much more widespread differential occupancy than previously suggested (Shen *et al*, 2012). Cell type-specific CTCF sites significantly overlapped with enhancers suggesting a role of these insulator sites in regulation of gene expression. On the other hand, it has been reported that CTCF-binding sites conserved between cell types represent an organizational pattern present in all cells, regardless of the developmental stage and tissue (Kim *et al*, 2007; Cuddapah *et al*, 2009) and it has been proposed that they delineate chromatin structures required for conserved genome functions that can be related to human diseases (Martin *et al*, 2011).

Several genome-scale mapping analysis of insulator protein-binding sites have been performed in flies and mammals (Kim *et al*, 2007; Bushey *et al*, 2009; Negre *et al*, 2010; Schwartz *et al*, 2012). Insulator proteins localize to thousands of sites characterized by conserved target sequences (Schwartz *et al*, 2012; Van Bortle *et al*, 2012). These observations and data obtained in the analysis of insulator protein recruitment in *Drosophila* in response to different stimuli, such as heat-shock response or ecdysone signaling (Wood *et al*, 2011) suggest that insulator activity must be controlled at the level of insulator protein binding to DNA, wherein differences in DNA motifs can influence protein occupancy levels and features of insulator function. It has been reported that distinct rules govern the binding of an

<sup>1</sup> Institute of Molecular Biology of Barcelona, IBMB-CSIC, and Institute for Research in Biomedicine IRB, Barcelona, Spain

<sup>2</sup> Bioinformatics and Biostatistics Unit, Institute for Research in Biomedicine IRB, Barcelona, Spain

\*Corresponding author. Tel: +34 93 4034962; Fax: +34 93 4034979; E-mail: mlebm@ibmb.csic.es

insulator protein to different classes of sites, which sometimes involve cooperation between several insulator proteins. It has been also shown that there are a number of CP190-binding sites that are not shared with any other known insulator protein, suggesting that there should exist more proteins that cooperate with CP190 to regulate insulator activity (Schwartz *et al*, 2012).

Here, we report on the identification of two novel DNA-binding proteins that mediate CP190 recruitment to chromatin and on the characterization of their role in insulator activity.

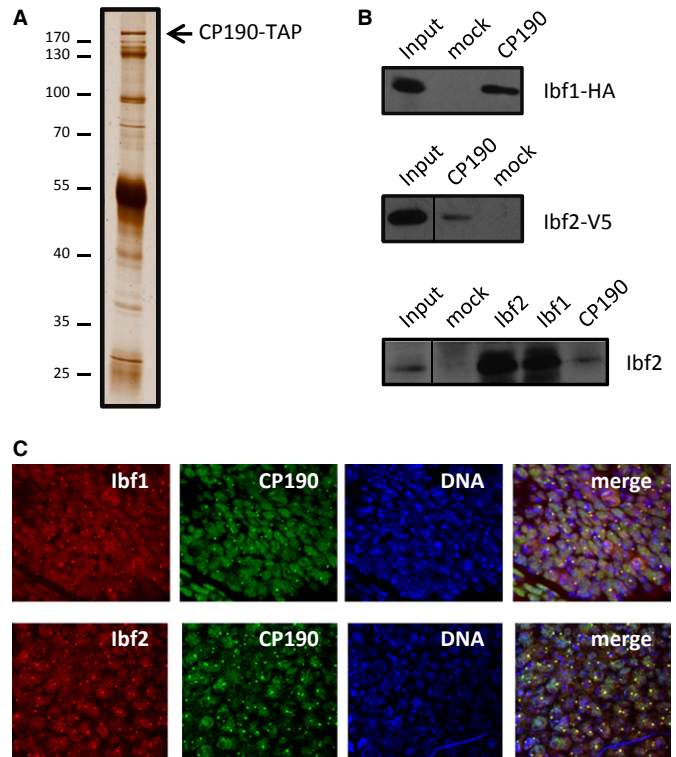
## Results

### Identification of two novel CP190-interacting proteins

Multiprotein complexes containing CP190 were purified to identify previously undescribed interacting proteins. To this end, a S2 *Drosophila* embryonic cell line expressing a CP190-TAP fusion protein was obtained. Nuclear extracts from that stable cell line were used to purify proteins that are associated with the CP190-TAP following a modified TAP procedure (see Materials and Methods). Eluates were analysed by SDS-polyacrylamide gradient gel electrophoresis which revealed multiple bands (Fig 1A). LC/MS analyses without prior electrophoretic separation were carried out and several different polypeptides that copurified with CP190-TAP in three independent purifications were identified (Table 1 and Supplementary Table S1). As expected, some of them corresponded to previously characterized CP190-associated proteins such as CP60 and the insulator factors CTCF, mod(mdg4)67.2 and Su(Hw) (Table 1). In addition to them, the products of two genes, CG8436 and CG9740 that consistently copurified with CP190-TAP were identified (Mascot scores of 500–250 and several identified peptides corresponding 30–50% sequence coverage). Since the molecular function of these proteins was not known we propose to name CG8436 Insulator binding factor 1 (Ibf1) and CG9740 Insulator binding factor 2 (Ibf2). A few other polypeptides were also identified but with lower scores.

We performed coimmunoprecipitation experiments to confirm the interaction between CP190, Ibf1 and Ibf2 and to this end we raised antibodies against CP190, Ibf1 or Ibf2 proteins (see Materials and Methods and Supplementary Fig S1). Epitope-tagged proteins, Ibf1-HA and Ibf2-V5, were transiently expressed in S2 cells and immunoprecipitation with  $\alpha$ CP190 antibodies resulted in coprecipitation of both Ibf1-HA and Ibf2-V5 (Fig 1B, upper and middle panels). No coimmunoprecipitation was observed when no antibodies were added. Moreover, immunoprecipitation assays of endogenous proteins showed that Ibf2 was significantly enriched in the immunoprecipitated samples with the different antibodies (Fig 1B, lower panel). This assay also showed that immunoprecipitation with  $\alpha$ Ibf1 results in quantitative coprecipitation of Ibf2 suggesting that both proteins form a stable complex in the cell.

Having identified Ibf1 and Ibf2 as CP190-interacting proteins we asked whether these proteins colocalize in the cell. Immunolocalization assays in third instar larvae imaginal discs showed that both proteins are distributed in a non-uniform pattern in nuclei of diploid cells where they overlap with CP190 at insulator body sites (Fig 1C).



**Figure 1. Ibf1 and Ibf2 associate with CP190.**

A Silver staining of 10% of the eluted proteins from CP190-TAP purification.  
B Western blot using  $\alpha$ HA (upper panel) or  $\alpha$ V5 (middle panel), and  $\alpha$ Ibf2 (lower panel) of nuclear extracts of either transiently transfected S2 cells with tagged-proteins (upper and middle panel) or S2 cells (lower panel) that were subjected to immunoprecipitation with the indicated antibodies ( $\alpha$ CP190,  $\alpha$ Ibf1,  $\alpha$ Ibf2) or no antibodies (mock). Input corresponds to 10% of the immunoprecipitated material.  
C Immunolocalization of CP190 (green) and Ibf1 (red, upper panel) or Ibf2 (red, lower panel) in imaginal disc cells. DNA is stained with DAPI (blue).

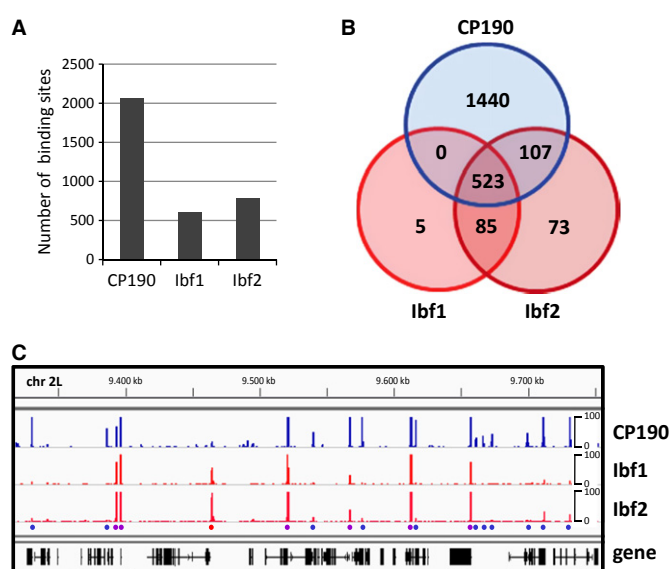
Source data are available online for this figure.

### Ibf1 and Ibf2 colocalize with CP190-binding sites throughout the genome

In order to obtain high-resolution information about the genome-wide chromatin association of Ibf1 and Ibf2 and their relationship with CP190 distribution pattern, we performed ChIP-seq analyses in S2 cells using  $\alpha$ CP190,  $\alpha$ Ibf1 and  $\alpha$ Ibf2 antibodies. We found 2,070 CP190 peaks, 613 Ibf1 peaks and 788 Ibf2 peaks (Fig 2A). Comparison of Ibf1 and Ibf2 ChIP-seq profiles indicates extensive overlap between them (Fig 2B, total number of binding sites and Fig 2C, binding profiles over a large region of chromosome 2L as an example). Indeed, more than 99% of Ibf1 overlap with Ibf2-binding sites and the higher signals observed in the Ibf2 ChIP-seq analysis are due, most likely, to technical differences in the immunoprecipitation efficiency (Fig 2C). These data along with the coimmunoprecipitation results reported above indicate that Ibf1 and Ibf2 form hetero-dimers that bind chromatin altogether. These analyses also showed that CP190 binds to most of the Ibf1 and Ibf2 sites (85% and 80% respectively) whereas only 25% of CP190 binding sites are also sites for Ibf1 and Ibf2 (Fig 2B and C). Therefore, these

**Table 1. CP190-interacting partners identified. List of identified CP190-interacting proteins ranked by the total number of peptides found by mass spectrometry in one representative experiment.**

Name	Peptides	Score	Coverage (%)	Function
CP190	39	2676	47.4	Insulator protein
CP60	17	1135	48.0	Microtubule-associated protein
Su(Hw)	10	547	13.0	Insulator protein
CG8436	9	619	49.2	Unknown
Mod(mdg4)67.2	7	502	17.7	Insulator protein
CG9740	5	341	51.8	Unknown
Pita	4	193	10.0	DNA binding protein
CTCF	2	199	6.0	Insulator protein

**Figure 2. Genome-wide Ibf1, Ibf2 and CP190 colocalization.**

- A A summary of the whole-genome peak analysis for CP190, Ibf1 and Ibf2. We defined peaks as locations within the enriched regions with a coverage difference between IP and input above 100 reads.
- B A Venn diagram indicating the number of binding sites that overlap between CP190, Ibf1 and Ibf2.
- C Representative ChIP-seq data for CP190, Ibf1 and Ibf2 over a 400-kb region of chromosome 2L. CP190 peaks are indicated by a blue circle, Ibf1/Ibf2 peaks are indicated by a red circle and overlapping peaks are indicated by purple circles.

results suggest that Ibf1 and Ibf2 participate in only a subclass of CP190 insulators.

We have also examined the location of Ibf1 and Ibf2 binding sites in relation to genetic features and we have analyzed the distribution based on whether sites are located either in intergenic regions, inside genes or overlapping the start or the end of the most proximal gene. As reported for other insulator proteins, Ibf1 and Ibf2 are preferentially located in intergenic regions (intergenic + promoter proximal in Supplementary Fig S2). Ibf1 and Ibf2 distribution patterns are essentially identical between them and similar to CP190 pattern although CP190 is more enriched at transcriptional start sites than the others (Supplementary Fig S2).

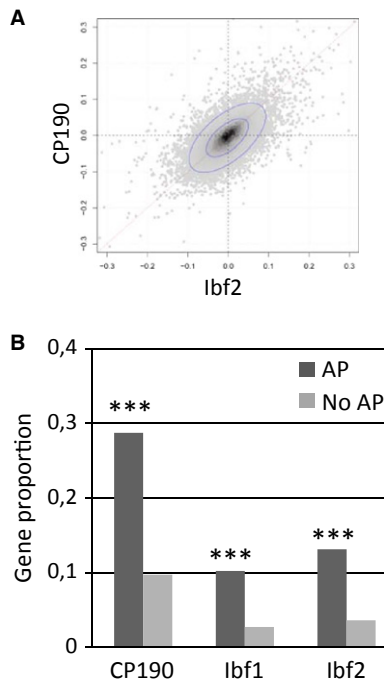
We next compared Ibf1 and Ibf2 genomic locations with that of the other insulator proteins taking advantage of the ChIP-chip/ChIP-seq results of the modENCODE database. More than 90% of CP190-binding sites from this study correspond to CP190 sites in the ChIP-chip assay of the modENCODE project being the much greater number of CP190-binding sites found by the modENCODE project likely the consequence of a difference in the levels of stringency applied (Supplementary Fig S3, compare upper and lower panels). Moreover, similar to the comparison reported above, most of Ibf1 and Ibf2 binding sites overlap with modENCODE CP190 sites (Supplementary Fig S4).

Comparison of Ibf1 and Ibf2 with the other insulator-binding proteins showed a stronger overlap between Ibf1/Ibf2 and CTCF peaks than either Su(Hw) or BEAF (Supplementary Fig S4). Recently it has been shown that CP190-binding sites can be found either alone or overlapping with different combinations of the insulator proteins Su(Hw), CTCF and BEAF (Schwartz *et al*, 2012). We found that Ibf1 and Ibf2 are present in only some of the co-binding groups reported by Schwartz *et al* (Supplementary Fig S5). Indeed, about 70% of co-bound CP190-CTCF sites, either containing Su(Hw) and mod(mdg4) or not, are also bound by Ibf1 and Ibf2 while these proteins can rarely be found in CP190-BEAF or CP190-Su(Hw)-mod(mdg4) groups. On the other hand, Ibf1 and Ibf2 can be found in some of the CP190 standalone sites but not in CTCF, Su(Hw) or BEAF standalone sites. Altogether these results indicate that Ibf1 and Ibf2 always associate with CP190 insulators which can also contain other insulator proteins, preferentially CTCF.

### Ibf1 and Ibf2 cooperate with CP190 to regulate gene expression

To characterize a role of Ibf1 and Ibf2 on gene expression and their relationship with CP190 we analyzed genomic changes in gene expression after depletion of these proteins in S2 cells. Since we could not obtain a significant knock-down after Ibf1 RNAi treatment we performed the analyses for CP190 and Ibf2 (see later in the text results indicating that Ibf2 mutation is likely to correspond to Ibf1/Ibf2 mutation regarding chromatin regulation). RNAi treatments decrease both CP190 and Ibf2 mRNA levels 4–5-fold and protein levels 2–3-fold (Supplementary Fig S6). As previously reported for CP190 and other insulator proteins (Bartkuhn *et al*, 2009; Schwartz *et al*, 2012), there is only a small number of genes that showed clear changes in their levels of expression after CP190 or Ibf2 knock-downs. We found both repressive and activating effects on gene



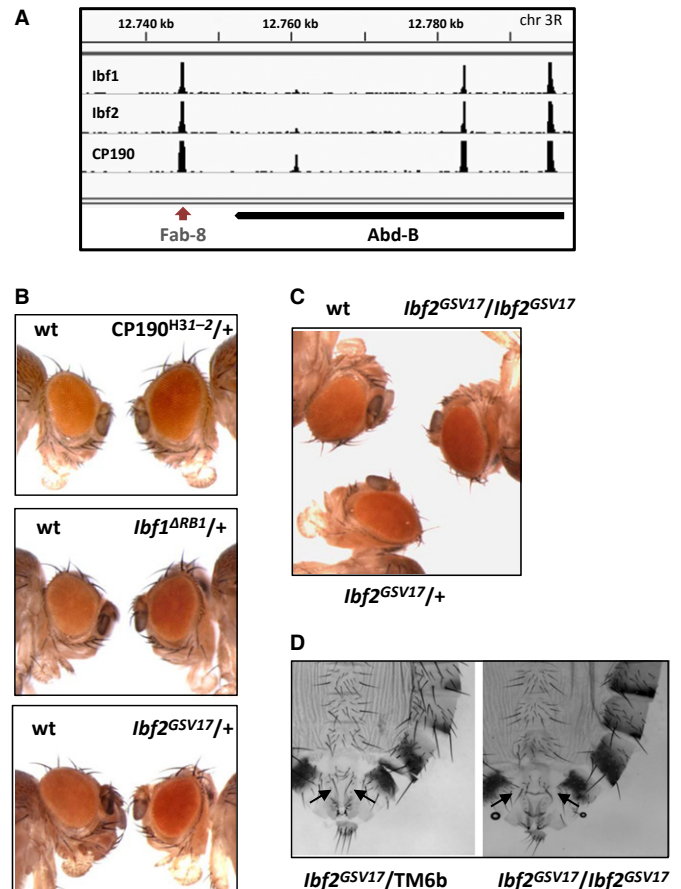


**Figure 3. Ibf1 and Ibf2 cooperate with CP190 to regulate gene expression.**

- A Log2 fold-changes in knockdown Ibf2 or CP190 conditions versus mock RNAi are plotted for visualization of the correlation between both conditions. Each dot represents a gene. Blue circles indicate 50% and 90% of total number of genes, respectively. The diagonal line corresponds to a perfect correlation of +1.
- B The graph shows proportion of genes containing CP190, Ibf1 or Ibf2 peaks. AP indicates genes with alternative promoters and No AP genes without alternative promoters. \*\*\* indicates that differences are statistically significant with  $P < 0.0001$  in Fisher exact tests.

expression. As shown in Fig 3A genes that change significantly their expression levels (those outside the blue circles) are mainly located in the upper right and lower left quadrants, meaning that they are genes that change in the same direction in both depletions. However, it must be noted that several of those differentially expressed genes do not fall close to the diagonal meaning that they are highly affected in one knockdown but not in the other. In order to evaluate the degree of co-regulation between Ibf2 and CP190 we plotted changes in expression for all the genes in the two knockdown conditions (Fig 3A) and we computed the Spearman's rank correlation between log2 fold changes. The Spearman's correlation value obtained was 0.588 ( $P$  value  $< 0.0001$ ), suggesting that these proteins cooperate in the regulation of gene expression.

Moreover, as it has been previously reported for vertebrate CTCF-binding sites (Kim *et al*, 2007), we have found that there is a significant concentration of Ibf1/Ibf2/CP190-binding sites at genes that display alternative promoter usage (Fig 3B). Indeed, the proportion of genes with different TSSs that contain peaks of these proteins is threefold compared with genes that do not contain alternative promoters suggesting that Ibf1, Ibf2 and CP190 can be functionally linked in the regulation of levels of alternative transcripts or their tissue specificity.



**Figure 4. Insulator activity of Ibf1 and Ibf2.**

- A ChIP-seq data for Ibf1, Ibf2 and CP190 on the *Abd-B* gene. Location of the Fab-8 boundary element is indicated.
- B Eye color of representative flies in enhancer-blocking assays using the transgenic line F8134.1 in wild-type and heterozygous *CP190*<sup>H31-2</sup>, *Ibf1*<sup>ΔRB1</sup> and *Ibf2*<sup>GSV17</sup> mutant backgrounds.
- C Eye color of representative flies in enhancer-blocking assays using the transgenic line F8109.6 in wild-type, heterozygous and homozygous *Ibf2*<sup>GSV17</sup> mutant backgrounds.
- D Abdomen cuticles of *Ibf2*<sup>GSV17</sup> heterozygous and homozygous females. Arrows indicate bristles in the A7 sternite showing different orientations, either toward the mid axis in heterozygous females (left panel) or towards the lateral sides in homozygous females (right panel).

### Ibf1 and Ibf2 are necessary for insulator activity

We next asked whether Ibf1 and Ibf2 play a role in insulator function and to answer this question we performed enhancer-blocking assays. The ChIP-seq assays reported above showed strong Ibf1 and Ibf2 binding to the Fab-8 boundary element of the BX-C (Fig 4A). Thus, to perform enhancer-blocking assays we obtained transgenic lines containing the Fab-8 boundary element (Barges *et al*, 2000) located between the white enhancer and the miniwhite reporter gene (see Materials and Methods). We obtained several lines showing enhancer-blocking activity of Fab-8 (see Materials and Methods, Fig 4 and Supplementary Fig S7).

On the other hand we obtained mutant alleles of Ibf1 and Ibf2 in a white background by mobilizing P elements present in either

CG8436/*Ibf1* or CG9740/*Ibf2* genomic regions (Supplementary Fig S8). The PBac[RB]CG8436<sup>e03576</sup> strain, which express normal levels of Ibf1 (data not shown), contains a P element inserted in the first intron of *Ibf1* close to the second exon of the gene. After mobilization of the transposon one of the lines obtained (*Ibf1*<sup>ARB1</sup>) harbors a deletion of DNA sequences from the second exon until the end of the *Ibf1* gene and the promoter and part of the coding sequence of the contiguous *VhaM8.9* gene. Therefore, this strain is expressing only the first exon of *Ibf1*, likely rendering the gene nonfunctional. However, it also corresponds to a *VhaM8.9* mutant allele, since the deletion includes the transcription start site and part of the coding sequences of this gene. Homozygotes *Ibf1*<sup>ARB1</sup> did not reach third instar larval stage due to, at least in part, *VhaM8.9* mutation. The p[GSV6]GS16482 strain contains a P element inserted in the coding sequence of the *Ibf2* gene which interrupts the Ibf2 protein at amino acid 22. After mobilization of this element we selected for lines (*Ibf2*<sup>GSV</sup>) that no longer express the *white* gene from the transposon sequences but still contain part of the transposon that renders the *Ibf2* gene nonfunctional (see Fig 6C, upper panel, Western blot assay showing absence of Ibf2 protein in *Ibf2*<sup>GSV17</sup> mutant background). Both, *Ibf2*<sup>GSV17</sup> homozygotes and *Ibf2*<sup>GSV17</sup>/Df(3R)ED5339 transheterozygotes showed reduced viability (60–70%), male sterility and many of the eclosed flies were short-lived.

We found a significant relief in enhancer-blocking using transgenic lines containing the Fab-8 boundary element in both *Ibf1* and *Ibf2* mutant backgrounds, since heterozygote *Ibf1*<sup>ARB1</sup> or *Ibf2*<sup>GSV17</sup> mutant flies showed a clear increase in the eye pigmentation similar to the one observed in the *CP190*<sup>H31-2</sup> mutant background (Fig 4B). Indeed, 66 out of 74 *Ibf1*<sup>ARB1</sup> heterozygote individuals and 149 of 152 *Ibf2*<sup>GSV17</sup> heterozygote individuals have darker eyes compared with wild-type flies. Similar effects were obtained using other independent Fab-8 transgenic lines and another *Ibf2* mutant background, *Ibf2*<sup>GSV20</sup> (Supplementary Fig S7). Moreover, a small but significant increase in eye color was seen in *Ibf2*<sup>GSV17</sup> homozygous condition (Fig 4C). Thus, we can conclude that Fab-8-mediated enhancer-blocking requires both Ibf1 and Ibf2.

Furthermore, *Ibf2*<sup>GSV17</sup> homozygous females exhibit a loss-of-function homeotic transformation similar to the one observed in *Fab-8*<sup>416</sup> flies, which carry a deletion of the Fab-8 boundary element (Barges et al, 2000). In adult mutant females the bristles in the A7 sternite have lost their orientation, which in wild-type is toward the mid axis, and resemble that normally found in more anterior sternites. Indeed, bristles in 70% of *Ibf2*<sup>GSV17</sup> homozygous females exhibit a vertical outward orientation indicating A7 to A6 transformation while only 14% of *Ibf2*<sup>GSV17</sup> heterozygous females show this homeotic transformation (Fig 4D). This homeotic transformation is also present at similar frequency in homozygotes for the P-element insertion in the original p[GSV6]GS16482 strain and in *Ibf2*<sup>GSV17</sup>/Df(3R)ED5339 transheterozygotes (75% and 72% respectively).

### Ibf1 and Ibf2 are related proteins that contain a BED finger DNA-binding domain

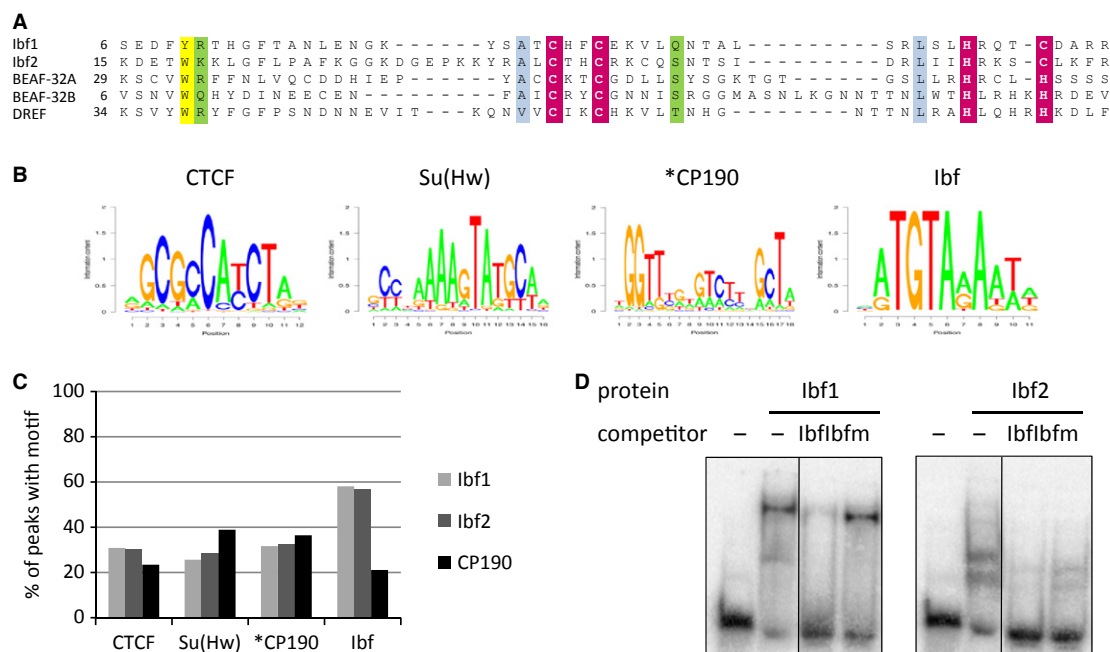
Ibf1 and Ibf2 are encoded by two genes, CG8436 and CG9740, located one next to the other in chromosome 3R at position 85D6 (Supplementary Fig S8). These two genes are transcribed in divergent orientations with TSS separated by 470 bp, giving rise to two polypeptides of 242 and 195 amino acids respectively. A BLAST

search of similar proteins in the *Drosophila melanogaster* databases performed with either Ibf1 or Ibf2 retrieved in both cases just one hit, which was either Ibf2 or Ibf1 respectively. Sequence alignment by ClustalW showed 18.7% of identity between both proteins with a highly conserved region in the N-terminal part of them (Supplementary Fig S9). A visual inspection of these regions showed a pattern of cysteines and histidines which could form a metal-chelating structure. Further analysis revealed that these proteins contain a BED finger, a DNA-binding domain that is present in the chromatin-boundary element binding proteins BEAF and DREF (Aravind, 2000). They share a signature of Cx<sub>2</sub>Cx<sub>n</sub>Hx<sub>3-5</sub>[H/C], with x<sub>n</sub> a variable spacer that is predicted to form a zinc finger (Fig 5A). BED proteins also contain a motif of two highly conserved aromatic residues separated by a polar residue which is partially conserved in Ibf1 and Ibf2 proteins. Besides, these proteins also possess a region enriched in basic residues located N-terminal to the BED domain, which in other proteins has been suggested to be involved in DNA contacts (Aravind, 2000).

The presence of DNA-binding ability in diverse BED domains suggests that Ibf1 and Ibf2 could bind directly to DNA. Therefore, we searched for conserved motifs in the ChIP-seq data using the rGADEM package. In Ibf1, Ibf2 and CP190 binding sequences we found not only the canonical CTCF and Su(Hw) motifs and a motif that has been reported to be related to CP190 binding, either directly or by means of an unknown protein (Schwartz et al, 2012), but also a new motif that we have named Ibf (Fig 5B). While CTCF, Su(Hw) and CP190 consensus motifs are present in similar percentages in either Ibf1, Ibf2 or CP190 binding sequences (Fig 5C), the Ibf motif was preferentially enriched in Ibf1 or Ibf2 binding sites (58–57%) compared to CP190-binding sites (21%). Moreover, this motif is present in only 7% of the CP190 peaks that do not contain Ibf1/Ibf2 whilst it is present in 50% of the CP190-binding sites that also bind Ibf1/Ibf2 (data not shown). Therefore, we hypothesized that this motif corresponds to the Ibf motif and we analyzed DNA binding *in vitro* by electrophoretic mobility shift assays (EMSA). We used a radioactively labeled oligonucleotide containing the Ibf motif and bacterially expressed purified Ibf1 and Ibf2 proteins (Fig 5D). These assays showed that both proteins bind specifically to the Ibf sequence since the retarded complex was eliminated when non labeled Ibf oligonucleotide was used as competitor but not when the competitor contains a mutated Ibf motif (see Materials and Methods for sequences). Thus, Ibf1 and Ibf2 bind directly to DNA suggesting that they could be involved in CP190 recruitment to chromatin at sites where they colocalize.

### Ibf1/Ibf2 complex mediate CP190 binding to chromatin

To further characterize the mechanisms involved in Ibf1 and Ibf2 insulator function we analyzed whether Ibf1, Ibf2 and CP190 binding to chromatin depend on each other. To this end we performed immunolocalization assays in mutant alleles using either the *Ibf2*<sup>GSV17</sup> mutant allele that corresponds to a null *Ibf2* mutation (see above in the text and Fig 6C) or the combination of *CP190*<sup>P1</sup>/*CP190*<sup>P11</sup> alleles that behaves as a null *CP190* mutation (Pai et al, 2004). These assays showed that the levels of Ibf1 on polytene chromosomes in the *Ibf2*<sup>GSV17</sup> mutant allele are dramatically reduced (Fig 6A). Since lack of Ibf1 binding in the Ibf2 mutant background could also reflect a contribution to synthesis and/or stability of the



**Figure 5. Characterization of Ibf1 and Ibf2 DNA-binding activities.**

A Alignment of Ibf1, Ibf2, BEAF and DREF BED finger domains. Numbers on the left side indicate the position of the amino acid in the corresponding protein. The shading is according to the BED finger domain consensus: the metal-chelating residues are shown in red shading whereas yellow, green and blue shading indicate aromatic, polar and hydrophobic residues respectively.

B The logo representation of sequence motifs characteristic of CTCF, Su(Hw), CP190 and Ibf binding sites defined by the rGADEM algorithm. \* in CP190-binding site indicates that the motif has been reported to be related to CP190 binding either directly or by means of an unknown protein.

C The presence of CTCF, Su(Hw), CP190 and Ibf motifs within Ibf1, Ibf2 and CP190 genomic binding sites.

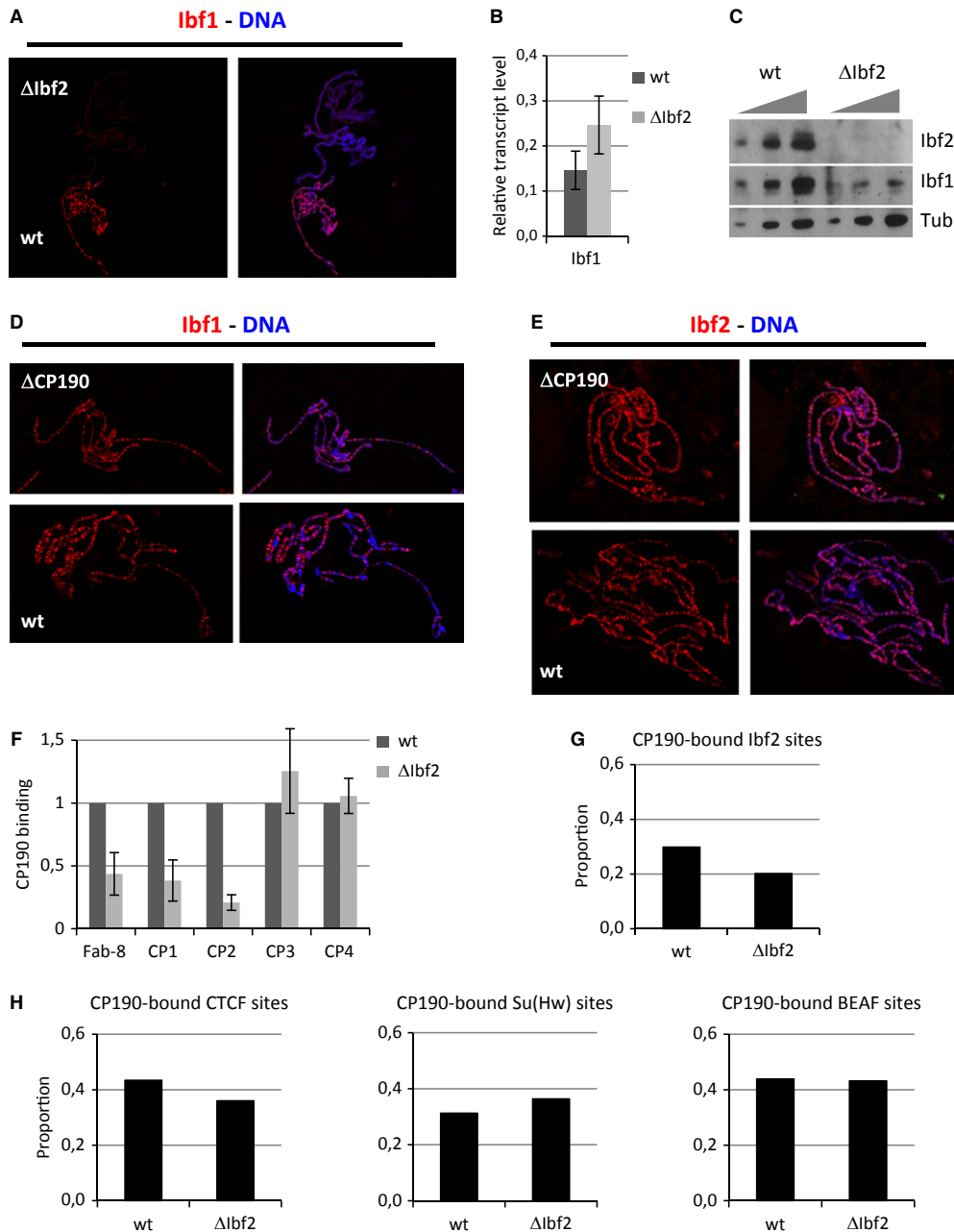
D EMSA analyses of Ibf1 and Ibf2 DNA-binding activity. The Ibf motif probe was incubated with bacterially expressed proteins. Where indicated, a competitor oligonucleotide containing either the Ibf motif (Ibf) or a mutated Ibf motif (Ibfm) was added.

Source data are available online for this figure.

protein we have analyzed Ibf1 mRNA and protein levels in Ibf2 mutant lines. While Ibf1 mRNA levels are slightly increased in Ibf2 mutant larvae (Fig 6B), protein levels are reduced by a factor between two and four (Fig 6C). These results strongly suggest destabilization of Ibf1 protein due, most likely, to its inability to associate with Ibf2 in the complex that binds chromatin. Similar results were reported for several components of the polycomb PRC2 complex when their binding to chromatin was impaired (Tan *et al*, 2007). On the contrary, both Ibf1 and Ibf2 appear normal in CP190 mutant polytene chromosomes (Fig 6D and Fig 6E respectively). Altogether these results indicate that Ibf1 needs Ibf2 to associate with chromatin whereas these proteins do not require CP190 to bind to chromatin.

On the other hand, CP190 appears to be present at normal levels in polytene chromosomes in the *Ibf2* mutant background (Supplementary Fig S10A). However, the results reported above show that the majority of CP190-binding sites do not correspond to *Ibf1* and *Ibf2* sites (Fig 2) and therefore shouldn't be affected in *Ibf1/Ibf2* mutant conditions rendering the analyses in polytene chromosomes not conclusive. To unambiguously determine whether CP190 binding depends on *Ibf1/Ibf2* we performed ChIP assays in wild-type and *Ibf2*<sup>GSV17</sup> mutant larvae that behave as an *Ibf1/Ibf2* null mutation regarding chromatin regulation, since we have found that *Ibf1*

is not able to bind chromatin in the absence of Ibf2 (Fig 6A). Our analyses showed a reduction in CP190 binding in the Ibf2 mutant background at the Fab-8 boundary element of the BX-C, where Ibf1 and Ibf2 colocalize with CP190 (Fig 6F). Moreover, our results also show a clear reduction in CP190 binding at other sites where the three proteins colocalize (CP1 and CP2 in Fig 6F) while CP190 remained associated to regions that do not correspond to Ibf1/Ibf2 binding sites (CP3 and CP4 in Fig 6F). In order to obtain genome-wide information about the effect of Ibf depletion on CP190 binding to chromatin we performed ChIP-seq analyses in wild-type and *Ibf2*<sup>GSV17</sup> mutant larvae using  $\alpha$ CP190 antibodies. These assays show that in wild-type larvae CP190 is bound to a fraction of the Ibf1/Ibf2 binding sites reported above (Fig 6G). This is in agreement with a previous report showing cell type-specific binding for another insulator protein (Shen *et al*, 2012). These assays also show that depletion of Ibf2 causes a reduction in the presence of CP190 on Ibf-binding sites (Fig 6G and see as an example Supplementary Fig S10, peak highlight with a red asterisk). Moreover, there are many CP190 peaks which, although still present in the mutant background, show a clear reduction in the levels of the protein (see for example Supplementary Fig S10, peak highlight with a blue asterisk). Altogether, these results indicate that Ibf proteins participate in CP190 recruitment to chromatin at many sites where they colocal-



**Figure 6. Ibf1/Ibf2 complex mediates CP190 binding to chromatin.**

A Patterns of immunolocalization of Ibf1 (red) in polytene chromosomes obtained from mutant *Ibf2*<sup>GSV17</sup> larvae that were mixed and squashed together with control wild-type larvae. DNA is stained with DAPI (blue).

B Quantitative RT-PCR analysis of the levels of Ibf1 in control wild-type and mutant *Ibf2*<sup>GSV17</sup> larvae. Relative Ibf1 mRNA expression was determined in relation to Rp49 expression. Data are presented as mean  $\pm$  s.d.

C Western blot analyses with  $\alpha Ibf1$  or  $\alpha Ibf2$  antibodies of protein extracts (twofold increasing amounts) from control wild-type and mutant *Ibf2*<sup>GSV17</sup> larvae. The signal obtained with  $\alpha$ Tubulin antibodies was used as a loading control.

D Patterns of immunolocalization of Ibf1 (red) in polytene chromosomes obtained from mutant *CP190*<sup>P1/P11</sup> and wild-type larvae. DNA is stained with DAPI (blue).

E Patterns of immunolocalization of Ibf2 (red) in polytene chromosomes obtained from mutant *CP190*<sup>P1/P11</sup> and wild-type larvae. DNA is stained with DAPI (blue).

F ChIP-qPCR using antibodies against CP190 and primers for the Fab-8 boundary element and a set of randomly chosen CP190-binding sites that are either Ibf1/Ibf2 binding sites (CP1 and CP2) or not (CP3 and CP4). Average enrichments (normalized to the input sample) are plotted as the ratio of precipitated DNA in wild-type or *Ibf2*<sup>GSV17</sup> mutant relative to wild-type larvae. Error bars show s.e.m. of three independent biological replicates.

G Proportion of CP190 peaks in either wild-type or mutant *Ibf2*<sup>GSV17</sup> larvae that overlap Ibf2 sites identified above (Fig 2).

H Proportion of CP190 peaks in either wild-type or mutant *Ibf2*<sup>GSV17</sup> larvae that overlap CTCF, Su(Hw) or BEAF sites (modENCODE \_283\_331 and \_274 files respectively).

Source data are available online for this figure.



ize. Nevertheless, these assays also show that there are some CP190-Ibf2 overlapping sites that are unaffected by Ibf2 depletion (see for example Supplementary Fig S10, peak highlight with a black asterisk) suggesting that CP190 binding to these sites depends on other insulator DNA-binding proteins. As reported above, our results point to a special participation of Ibf1/Ibf2 complex in CP190 insulators that also contain CTCF. In order to characterize this further we analyzed the proportion of CP190-binding sites in either wild-type or mutant *Ibf2*<sup>GSV17</sup> larvae that overlap CTCF, Su(Hw) or BEAF sites. These analyses reveal that CP190 sites that are lost in the Ibf2 mutant condition overlap significantly with CTCF sites while CP190 sites that are maintained in the absence of Ibf2 more often correspond to Su(Hw) sites (Fig 6H).

Overall, we can conclude that the Ibf1/Ibf2 complex mediates CP190 recruitment to chromatin at sites where they colocalize and that these proteins are required for insulator function.

## Discussion

In this study we have identified Ibf1 and Ibf2 as two novel CP190-interacting proteins and we have shown that these proteins are able to bind DNA with affinity and specificity. The mechanistic interdependencies between CP190 and sequence-specific DNA-binding proteins such as Su(Hw), CTCF and BEAF, although extensively analyzed in immunolocalization and ChIP assays, are still not fully understood. It has been shown that CP190 depends on Su(Hw) to associate with the gypsy insulator but this dependence is not clear for endogenous Su(Hw)-dependent insulators (Pai *et al*, 2004). CP190 and CTCF cooperate with each other to bind chromatin but different studies showed different degrees of interdependence (Gerasimova *et al*, 2007; Mohan *et al*, 2007; Wood *et al*, 2011). Recently, a study taking into account quantitative aspects of binding revealed several classes of binding sites occupied by specific combinations of insulator proteins and proposed the existence of still unknown players (Schwartz *et al*, 2012). Our results demonstrate a novel pathway of CP190 recruitment to chromatin which involves binding to a complex formed by two previously uncharacterized proteins, Ibf1 and Ibf2. These findings add more complexity to insulators and could explain previously controversial observations on the mechanistic interdependencies between CP190 and the previously characterized DNA-binding proteins CTCF, Su(Hw) and BEAF. Nevertheless, there are still several CP190-binding sites that neither colocalize with previously known insulator proteins nor with Ibf1 and Ibf2 suggesting that there may be additional unknown DNA-binding proteins involved in insulator function in *Drosophila*.

In this study we have shown that Ibf1 and Ibf2 contain a BED finger which is a DNA-binding domain that is also present in the insulator protein BEAF and several transposases distributed over the eukaryotic crown group. This suggests that in *Drosophila* the BED finger has been recruited for cellular functions from transposases to regulate his own chromatin structure. It has been shown that BEAF insulator protein separates close head-to-head genes with different patterns of transcriptional regulation (Yang *et al*, 2012). On the other hand, previously reported studies showed concentration of CTCF-binding sites at genes that contain alternative promoters in vertebrates (Kim *et al*, 2007) and the correlation of CTCF binding to the protocadherin  $\alpha$  gene with alternative isoform expression (Mona-

han *et al*, 2012). Our results show that Ibf1, Ibf2 and CP190-binding sites are clearly enriched at genes with alternative promoters suggesting a role for these proteins in the selection of promoters in distinct cell types along development.

Recent data indicate that chromatin insulators outgrow the classical barrier and enhancer-blocking roles that defined these elements; they appear to be involved in mediating long-range inter- and intra-chromosomal arrangements that can direct the nuclear co-localization of specific sequences playing a crucial role in the regulation of gene expression. The insulator proteins CP190 in *Drosophila* and CTCF-cohesin in mammals have been shown to mediate protein-protein interactions resulting in the organization of the chromatin into loops. It is tempting to speculate that Ibf1 and Ibf2, besides recruiting CP190 to chromatin, can help this protein in the establishment or maintaining of the interactions between different insulator elements; we have shown that both, Ibf1 and Ibf2, are able to bind specifically to the same DNA motif but the results of genome-wide analyses presented here strongly suggest that they do not compete for binding to DNA. Ibf1 and Ibf2 are similar proteins encoded by two genes that are located one next to the other in the genome suggesting that they arose from a gene duplication event. Our analyses show that both proteins are required for insulator activity since mutation of any of them has an effect in enhancer-blocking assays and our results indicate that Ibf1 and Ibf2 form a stable complex in the cell. It has been shown that duplication of genes coding for homodimers is frequently accompanied by conservation of protein interactions permitting functional diversification (Pereira-Leal *et al*, 2007). Ibf1 and Ibf2 contain, besides the N-terminal DNA-binding domain, a conserved region at the C-terminal part that could account for Ibf1-Ibf2 protein interactions. Although further investigation will be needed to characterize whether both, Ibf1 and Ibf2, or which one of them, support chromatin contacts and/or protein-protein interactions with CP190, our results clearly show that Ibf1/Ibf2 complex binds to chromatin and to CP190 and they suggest that this complex might be able to interact with two different insulator elements contributing to the loop organization of the chromatin.

## Materials and Methods

### Fly stocks

p[GSV]GS16483 was obtained from DGRC (Indiana University), pBac[RB]CG8436<sup>e03576</sup> from the Exelixis Collection (Harvard Medical School) and Df[3R]ED5339 was obtained from Bloomington *Drosophila* Stock Center (Indiana University). CP190<sup>P1</sup> and CP190<sup>P11</sup> were obtained from V. Corces (Pai *et al*, 2004).

*Ibf2* mutant lines were obtained by mobilization of the P element from p[GSV]GS16483 followed by selection of lines that either kept part of the transposon or delete enough DNA sequences to also affect Ibf1 gene. We obtained 40 white minus lines, among them, none with a lesion affecting both genes and 13 lines that kept parts of the transposon. The integrity of the genomic sequences on both sides of the insert in lines *Ibf2*<sup>GSV17</sup> and *Ibf2*<sup>GSV20</sup> was checked by PCR (data not shown).

The Fab-8 sequence (Barges *et al*, 2000) was obtained by PCR amplification of a DNA fragment from 3R:12744472 to 3R:12745518 using the following primers: 5'-CTTTGACGAGTTTCCAAGC-3' and

5'-GGGCTACCAGTGCCTGGCC-3'. The PCR product was inserted in a pCasper between the enhancer of the *white* gene and a mini-white reporter gene, as described in (Perez-Lluch *et al*, 2008). When this construct was introduced into flies 60% of the transgenic lines obtained (five out of eight) showed reduced eye pigmentation ranging from yellow to dark orange. Lines with light eyes were used in enhancer-blocking assays and the eye color was determined 24 h after eclosion for control and mutant backgrounds.

### Antibodies

Rat  $\alpha$ -Ibf-1 and  $\alpha$ -Ibf-2 polyclonal antibodies were raised against bacterially expressed recombinant proteins containing the full-length Ibf1 and Ibf2 and were used at a 1:2,500 dilution for Western blot. Rabbit  $\alpha$ -CP190 polyclonal antibodies were raised against a bacterially expressed recombinant protein containing amino acids 125–1096 of CP190 and was used at 1:4,000 dilution for Western blot. Rat  $\alpha$ -HA was purchased from Roche (12CA5), mouse  $\alpha$ V5 was purchased from Invitrogen (R960-25), rabbit  $\alpha$ -GFP was purchased from Invitrogen (A-6455), and mouse  $\alpha$ - $\beta$ tubulin was purchased from Sigma (A2547).

### Purification and characterization of multiprotein complexes containing CP190

Purification of multiprotein complexes was performed from stable S2 cell lines expressing CP190-TAP. Full-length CP190 was TAP-tagged at the C-terminus using plasmid pMK33-CTAP where expression of the TAP-tagged protein is driven by a metallothionein inducible promoter, and stable S2 lines obtained according to (Veraksa *et al*, 2005). Expression of CP190-TAP was induced by overnight treatment with 0.07 mM CuSO<sub>4</sub>. Crude nuclear extracts were prepared according to (Dignam *et al*, 1983) and precleared by incubation with BSA-Dynabeads for 2 h at 4°C. Extracts were, then, applied to IgG-Dynabeads (Invitrogen) and incubated overnight at 4°C. After incubation, beads were extensively washed and eluted with 25 mM Tris-HCl (pH 6.8), 5% glycerol, 1% SDS. Eluted material was analysed by SDS-PAGE. A single band from the stacking gel was cut and analyzed by standard LC/MS, performed in the Proteomics Unit of the “Institut de Recerca de la Vall d’Hebron” (Barcelona).

### Coimmunoprecipitation experiments

Assays were performed with extracts prepared from S2 cells or transiently transfected S2 cells with Ibf1-HA (LD32555, obtained from DGRC) or Ibf2-V5 (full-length Ibf2 was cloned into a pAc5.1). Cells were lysed with 1% Triton X-100, 300 mM NaCl, 50 mM Tris pH 8, 1 mM EDTA, 1 mM DTT. Incubation with  $\alpha$ -CP190,  $\alpha$ -Ibf1,  $\alpha$ -Ibf2 or no antibodies was performed at 4°C for 2 h. After incubation with Protein A/G Agarose (Santa Cruz Biotech), beads were pelleted by centrifugation, washed and analysed by Western blot.

### Immunostaining experiments

Immunostaining experiments were performed in polytene chromosomes and imaginal discs of third instar larvae with  $\alpha$ -Ibf1 (1:400),  $\alpha$ -Ibf2 (1:500) and  $\alpha$ -CP190 (1:1,000) according to (Font-Burgada *et al*, 2008). For visualisation, slides were mounted in Mowiol

(Calbiochem-Novabiochem) containing 0.2 ng/ml DAPI (Sigma) and visualised in a Nikon Eclipse E-800 inverted microscope or in a confocal Leica TCS SP2-AOBS microscope.

### ChIP experiments

For ChIP, chromatin was prepared according to (Lloret-Llinares *et al*, 2012) from cultured S2 cells and from third instar larvae according to (Carbonell *et al*, 2013). Immunoprecipitations were basically performed as described in (Orlando *et al*, 1997) using rat  $\alpha$ -Ibf1,  $\alpha$ -Ibf2, and rabbit  $\alpha$ -CP190 polyclonal antibodies. For ChIP-qPCR, triplicates from three independent biological replicates were analyzed (see Supplementary Table S2 for primers used in these experiments).

For ChIP-seq, 10 ng of DNA, quantified by Qubit dsDNA HS Assay Kit (Invitrogen) were used for library preparation. End-repair, adenylation, ligation of adapters and PCR enrichment for 18 cycles was performed using TruSeqRNA Sample Prep Kit (Illumina) according to manufacturer's recommendations. Purified libraries were quantified by Qubit dsDNA HS Assay Kit (Invitrogen) and size distribution was evaluated using Bioanalyzer DNA 1000 assay (Agilent). Single-end sequencing of 50 nucleotides was performed on Solexa/Illumina. See Supplemental experimental procedures for details on data processing and quality control, sequence alignment, peak calling and downstream analyses.

### Expression profiling

S2 cells ( $2.5 \times 10^6$ ) were transfected with 15  $\mu$ g dsRNA. As a mock RNAi treatment cells were transfected with LacZ dsRNA (see Supplementary Table S2 for primers used to produce Ibf1, Ibf2, CP190 and LacZ dsRNAs). The transfection was repeated after 48 h and 72 h and cells collected for RNA extraction at 96 h, using Rneasy mini kit (Qiagen). Duplicates were processed for each RNAi condition.

cDNA library preparation and amplification were performed from 25 ng total RNA using WTA2 (Sigma-Aldrich) with 19 cycles of amplification. 8  $\mu$ g cDNA was subsequently fragmented by DNaseI and biotinylated by terminal transferase obtained from GeneChip Mapping 10Kv2 Assay Kit (Affymetrix). Hybridization mixture was prepared according to Affymetrix protocol. Each sample target was hybridized to a GeneChip Drosophila Genome 2.0 Array (Affymetrix). The arrays were scanned in a GeneChip Scanner 3000 (Affymetrix). CEL files were generated from DAT files using GCOS software (Affymetrix). To generate the log<sub>2</sub> expression estimates, overall array intensity was normalized between arrays and the probe intensity of all probes in a probeset summarized to a single value using RMA (Robust Multichip Average) algorithm (Irizarry *et al*, 2003).

To measure the degree of co-regulation between the mutants, we computed the pair-wise Spearman's rank correlation between samples.

### Electrophoretic mobility shift assays

DNA probes were obtained by annealing oligonucleotides for the Ibf motif: 5'-TCTGTGTTTTGACTATTTTACATTTGACCATGGAT-3' and 5'-TCTGATCCATGGTCAAATGTAAATAGTCAAAACA-3' and for a mutated Ibf motif: 5'-TCTGTGTTTTGACTGGCCTTCTGTTGAC

CATGGAT-3' and 5'-TCTGATCCATGGTCAACAGAAGGCCAGTCAAA ACA-3' in Tris pH 7.5 10 mM and MgCl<sub>2</sub> 10 mM. Annealed oligonucleotides were end-labeled using [ $\alpha$ -<sup>32</sup>P]dATP and Klenow fragment (Fermentas). Recombinant full-length Ibf1 or Ibf2 was added to the purified probe in a 10  $\mu$ l final volume of binding reaction containing 20% Glycerol, 50 mM Hepes pH 7.8, 120 mM KCl, 1  $\mu$ g/ $\mu$ l BSA, 0.01  $\mu$ g/ $\mu$ l ssDNA and 2 mM DTT. For DNA competition analysis, binding was competed using 60-fold excess of either Ibf probe or mutated Ibf probe. Binding reactions were incubated on ice for 15 min and then resolved on 8% non-denaturing polyacrylamide gels. Gels were dried and subjected to autoradiography.

### Accession numbers

ChIP-seq and expression profiling data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE47559.

**Supplementary information** for this article is available online: <http://emboj.embopress.org>

### Acknowledgments

We thank Fernando Azorín, Josep Casacuberta, Jordi Casanova and Joan Font-Burgada for helpful discussions and comments on the manuscript. We also thank Herbert Auer for technical support with genome-wide experiments and Estefanía Freire and Alicia Vera for technical assistance. We are thankful to Victor G. Corces, Bloomington Stock Center, DGRC and Harvard Medical School for providing fly stocks and clones. S. C. and U. F. acknowledge receipt of doctoral fellowships from MEC and CSIC respectively. Funding was from the Spanish Ministerio de Ciencia e Innovación (BFU2009-10983 and CSD2006-49) and the Generalitat de Catalunya (SGR2009-1023).

### Author contributions

Conceived and designed the experiments: SC, MLE. Performed the experiments: SC, UF, MLE. Analyzed the data: SC, UF, OR, EP, MLE. Wrote the paper: SC, MLE.

### Conflict of interest

The authors declare that they have no conflict of interest.

## References

- Ahanger SH, Shouche YS, Mishra RK (2013) Functional sub-division of the *Drosophila* genome via chromatin looping: the emerging importance of CP190. *Nucleus* 4: 115–122
- Aravind L (2000) The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. *Trends Biochem Sci* 25: 421–423
- Barges S, Mihaly J, Galloni M, Hagstrom K, Muller M, Shanower G, Schedl P, Gyurkovics H, Karch F (2000) The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. *Development* 127: 779–790
- Bartkuhn M, Straub T, Herold M, Herrmann M, Rathke C, Saumweber H, Gilfillan GD, Becker PB, Renkawitz R (2009) Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J* 28: 877–888
- Belozero V, Majumder P, Shen P, Cai HN (2003) A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of *Drosophila*. *EMBO J* 22: 3113–3121
- Bushey AM, Ramos E, Corces VG (2009) Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev* 23: 1338–1350
- Carbonell A, Mazo A, Serras F, Corominas M (2013) Ash2 acts as an ecdysone receptor coactivator by stabilizing the histone methyltransferase Trx. *Mol Biol Cell* 24: 361–372
- Chopra VS, Cande J, Hong JW, Levine M (2009) Stalled Hox promoters as chromosomal boundaries. *Genes Dev* 23: 1505–1509
- Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 19: 24–32
- Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* 11: 1475–1489
- Ebersole T, Kim JH, Samoshkin A, Kouprina N, Pavlicek A, White RJ, Larionov V (2011) tRNA genes protect a reporter gene from epigenetic silencing in mouse cells. *Cell Cycle* 10: 2779–2791
- Font-Burgada J, Rossell D, Auer H, Azorin F (2008) *Drosophila* HP1c isoform interacts with the zinc-finger proteins WOC and Relative-of-WOC to regulate gene expression. *Genes Dev* 22: 3007–3023
- Gerasimova TI, Lei EP, Bushey AM, Corces VG (2007) Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Mol Cell* 28: 761–772
- Gurudatta BV, Corces VG (2009) Chromatin insulators: lessons from the fly. *Brief Funct Genomic Proteomic* 8: 276–282
- Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31: e15
- Karch F, Galloni M, Sipos L, Gausz J, Gyurkovics H, Schedl P (1994) Mcp and Fab-7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of *Drosophila melanogaster*. *Nucleic Acids Res* 22: 3138–3146
- Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanov VV, Ren B (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128: 1231–1245
- Lloret-Llinares M, Perez-Lluch S, Rossell D, Moran T, Ponsa-Cobas J, Auer H, Corominas M, Azorin F (2012) dKDM5/LID regulates H3K4me3 dynamics at the transcription-start site (TSS) of actively transcribed developmental genes. *Nucleic Acids Res* 40: 9493–9505
- Lunyak VV, Prefontaine GG, Nunez E, Cramer T, Ju BG, Ohgi KA, Hutt K, Roy R, Garcia-Diaz A, Zhu X, Yung Y, Montoliu L, Glass CK, Rosenfeld MG (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* 317: 248–251
- Maeda RK, Karch F (2007) Making connections: boundaries and insulators in *Drosophila*. *Curr Opin Genet Dev* 17: 394–399
- Martin D, Pantoja C, Fernandez Minan A, Valdes-Quezada C, Molto E, Matesanz F, Bogdanovic O, de la Calle-Mustienes E, Dominguez O, Taher L, Furlan-Magaril M, Alcina A, Canon S, Fedetz M, Blasco MA, Pereira PS, Ovcharenko I, Recillas-Targa F, Montoliu L, Manzanera M et al (2011) Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. *Nat Struct Mol Biol* 18: 708–714
- Mohan M, Bartkuhn M, Herold M, Philippen A, Heintz N, Bardenhagen I, Leers J, White RAH, Renkawitz-Pohl R, Saumweber H, Renkawitz R (2007) The

- Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J* 26: 4203–4214
- Monahan K, Rudnick ND, Kehayova PD, Pauli F, Newberry KM, Myers RM, Maniatis T (2012) Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin- $\alpha$  gene expression. *Proc Natl Acad Sci USA* 109: 9125–9130
- Negre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, Feng X, Ahmad K, Russell S, White RA, Stein L, Henikoff S, Kellis M, White KP (2010) A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet* 6: e1000814
- Noma K, Cam HP, Maraia RJ, Grewal SI (2006) A role for TFIIIC transcription factor complex in genome organization. *Cell* 125: 859–872
- Oki M, Kamakaka RT (2005) Barrier function at HMR. *Mol Cell* 19: 707–716
- Orlando V, Strutt H, Paro R (1997) Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* 11: 205–214
- Pai C-Y, Lei EP, Ghosh D, Corces VG (2004) The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol Cell* 16: 737–748
- Pereira-Leal JB, Levy ED, Kamp C, Teichmann SA (2007) Evolution of protein complexes by duplication of homomeric interactions. *Genome Biol* 8: R51
- Perez-Lluch S, Cuartero S, Azorin F, Espinas ML (2008) Characterization of new regulatory elements within the Drosophila bithorax complex. *Nucleic Acids Res* 36: 6926–6933
- Raab JR, Chiu J, Zhu J, Katzman S, Kurukuti S, Wade PA, Haussler D, Kamakaka RT (2012) Human tRNA genes function as chromatin insulators. *EMBO J* 31: 330–350
- Schwartz YB, Linder-Basso D, Kharchenko PV, Tolstorukov MY, Kim M, Li HB, Gorchakov AA, Minoda A, Shanower G, Alekseyenko AA, Riddle NC, Jung YL, Gu T, Plachetka A, Elgin SC, Kuroda MI, Park PJ, Savitsky M, Karpen GH, Pirrotta V (2012) Nature and function of insulator protein binding sites in the Drosophila genome. *Genome Res* 22: 2188–2198
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanov VV, Ren B (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature* 488: 116–120
- Tan J, Yang X, Zhuang L, Jiang X, Chen W, Lee PL, Karuturi RK, Tan PB, Liu ET, Yu Q (2007) Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev* 21: 1050–1063
- Valenzuela L, Kamakaka RT (2006) Chromatin insulators. *Annu Rev Genet* 40: 107–138
- Van Bortle K, Corces VG (2013) The role of chromatin insulators in nuclear architecture and genome function. *Curr Opin Genet Dev* 23: 212–218
- Van Bortle K, Ramos E, Takenaka N, Yang J, Wahi JE, Corces VG (2012) Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res* 22: 2176–2187
- Veraksa A, Bauer A, Artavanis-Tsakonas S (2005) Analyzing protein complexes in Drosophila with tandem affinity purification-mass spectrometry. *Dev Dyn* 232: 827–834
- Wallace JA, Felsenfeld G (2007) We gather together: insulators and genome organization. *Curr Opin Genet Dev* 17: 400–407
- West AG, Gaszner M, Felsenfeld G (2002) Insulators: many functions, many mechanisms. *Genes Dev* 16: 271–288
- Wood AM, Van Bortle K, Ramos E, Takenaka N, Rohrbaugh M, Jones BC, Jones KC, Corces VG (2011) Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. *Mol Cell* 44: 29–38
- Yang J, Ramos E, Corces VG (2012) The BEAF-32 insulator coordinates genome organization and function during the evolution of Drosophila species. *Genome Res* 22: 2199–2207





